Psychological Adaptation to Counter-stereotypical Diversity

Ekaterina Damer

Department of Psychology

Faculty of Science

University of Sheffield

A thesis submitted for the degree of

Doctor of Philosophy

April 2018

Dedicated to Irina, Ingo, Eduard, and Phelim.

Danke für eure Unterstützung, ihr seid unschätzbar.

Я вас очень люблю.

Acknowledgements

I feel humbled by the amount of support I've received by my family, friends, and colleagues over the years.

Mum, you have always cherished education and personal growth, supporting me emotionally, intellectually, and financially in every way you could. Ingo, you've been a bastion of calm in the stormiest of times; your strength of character and attention to detail inspire me. Eduard, your courage and compassion have left a mark on me and I'm grateful to you for that. Phelim, my love, you have been invaluable to me in more ways than I can fathom. Your unconditional support, patience and kindness, healthy scientific scepticism, as well as humour, openness, and curiosity have shaped me forever. Together, you're my heroes.

My dear mentors Rich Crisp and Tom Webb, you could not have been a more complementary and intellectually stimulating supervisory team. You've taught me both to think big and to pay attention to nuance; to be rigorous, but not rigid; and to be optimistic yet level-headed. Thank you for being so reliable, thoughtful, and kind in this thrilling journey towards becoming an independent scientist.

I would like to thank my colleagues and collaborators——especially Marcel Wiechmann, Justina Mischewski, Patrick Forscher, Paul Maher, Tomas Folke, Laura Di Bella, Caro Verfürth, Lauren Aguilar, Jens Hellmann, and many more——for being inspiring freethinkers who care about the state of the world and want it improved. Our conversations and debates have been absorbing and fascinating; I have learned a lot from you.

A standing ovation goes to my team at Prolific——Phelim, Jim, Andrew, Stephen, Kendra, and Zak——for your trust and dedication. In the months nearing thesis completion I've had to exclusively focus on my academic work, but in my absence you've kept the business not just running, but also thriving. Thank you for having my back. I'm excited about what's coming next.

A precious thought goes to my lovely friends, old and new: Sina Kraus, Sonja Monsees, Edisa Korugic, Björn Freter, Agne Milukaite & Peter Ebsen, Nicole Miranda & Chris Birkl, and Harriet Baird. I deeply appreciate your friendship. Thank you for keeping me sane and for being absolutely great craic!

Finally, huge thanks go to my Department at the University of Sheffield for financial and logistical support provided, especially to Josie Cassidy, Andy Bassett, Abbie McGregor, and Paul Overton. When the opportunity presented itself to visit Stanford's Psychology Department for six

months in 2016, you've generously provided financial support, going out of your way to help me make the research visit happen. I'm also indebted to my fellow postgraduate students and the Psychology faculty members for thought-provoking discussions and great times together.

I've come a long way from Russia via Germany to the UK, and I've seen you and many others shatter stereotypes along the way. I'm delighted to have shared the journey with you all, and I hope that we'll continue forging our diverse and unique trajectories.

Abstract

Social and cultural diversity are globally increasing at an unprecedented pace. The implications of this increase for individuals and societies can vary: Benefits such as cognitive flexibility and creativity may ensue when groups manage to cooperate and integrate, but stress and conflict may follow when groups are segregated and marginalised. This thesis focuses on a form of diversity that challenges traditional stereotypes (e.g., a female entering a male-dominated profession) and is thus termed counter-stereotypical diversity. The aim is to empirically and theoretically explore how people psychologically adapt to counter-stereotypical diversity. One primary prediction was derived from the literature, which is that exposure to exemplars of counter-stereotypical diversity (termed counter-stereotypes, CSTs) can boost cognitive flexibility, and this was tested across 12 experiments (reported in Chapters 4, 5, and 6). Various secondary predictions were also tested, for example the role of need for cognition in moderating the effects of exposure to CSTs on cognitive reflection (Chapter 5), and the longitudinal effects of exposure to CSTs on cognitive flexibility and intergroup bias (Chapter 6). CSTs were conceptualised as a special case of expectancy violations, and Chapter 7 theorised that they can be followed by three types of responses: (1) indifference, (2) threat (and defensiveness), and/or (3) challenge (and open-mindedness). Overall, this thesis improves our understanding of how people psychologically adapt to counter-stereotypical diversity.

Contents

List of Figures

Chapter 1

Introduction

1.1 Rising Social and Cultural Diversity

Social and cultural diversity are globally increasing at an unprecedented pace. An estimated 258 million people live in a country other than their country of birth (Kobler, 2017). Europe alone is home to almost 100 distinct ethnic groups——twice as many as nation-states (Pan, Pfeil & Videsott, 2016)——who speak over 80 different languages (Eurobarometer, 2012). In some countries, minority groups are growing so quickly relative to the majority group that a "majority-minority crossover" will likely occur by 2050 in the US, and by 2070 in the UK, meaning that their populations will consist of more people of colour than Whites (Coleman, 2010; Ortman & Guarneri, 2009). The UK's immigrant population is one of the most diverse in the world, tying Denmark for the highest diversity score, which measures the distribution of

[1]The top three origin countries of immigrants living in the UK are India (780,000), Poland (700,000) and Pakistan (540,000)

immigrants from different origin countries (Pew Research Center, 2016).[1]

In addition, many people are exposed to new forms of diversity via the media. It is estimated that, in 2016, the average person spent 7.6 hours a day interacting with traditional or digital media (including printed newspapers, magazines, broadcast television, radio, and online media) as a way of learning news and engaging socially (Zenith Media, 2017). Despite concerns that "news-filtering algorithms narrow what we know" (Pariser, 2015, p. 1), recent evidence from the 2017 Reuters Institute Digital News Report suggests that social media use, for example, is linked to increased exposure to politically diverse content (Fletcher & Nielsen, 2017). These media trends underscore the importance of understanding more deeply how people respond to different types of offline and online information. In line with this view, this thesis adopts a methodological approach that exposes people to information about diverse individuals in order to examine whether such exposure may impact broader psychological functioning.

Taken together, it seems uncontroversial that social and cultural diversity have become defining and ubiquitous characteristics of many societies today. In short, our world seems to be turning into a "global village" in the sense that diverse people around the globe——irrespective of their exact location——are connected with other parts of the world through the media, thus giving rise to increasingly intertwined human dynamics (Martens, Dreher & Gaston, 2010). The consequences of increasing social and cultural diversity can vary: While contact and cooperation may sometimes improve intergroup relations (Allport, 1954, 1979), foster perspective taking skills (Hoever, van Knippenberg, van Ginkel & Barkema, 2012; Sommers, 2006), and

increase cognitive flexibility and creativity (Crisp & Turner, 2011), diversity may at other times lead to conflict (Pelled, Eisenhardt & Xin, 1999), acculturative stress (Berry, 1970; Berry, 1997), and social isolation (Putnam, 2007). The next section reflects on why it is in the interest of the global community (including scientists, educators, politicians, and business leaders) to effectively manage our increasingly diverse societies.

1.2 The Importance of Managing Adaptation to Rising Diversity Effectively

Humans have a basic preference for homogeneity, stability, and structure (Abrams & Hogg, 1988; Caporael, 1997; Neuberg & Newsom, 1993), which can manifest in a desire to separate into ingroups versus outgroups, or "us" versus "them" (Crisp, 2015; Tajfel, Billig, Bundy & Flament, 1971; Tajfel & Turner, 1979). This fundamental insight poses an important and challenging question: Given this need for coherence, how can people adapt to a multicultural future? Some have argued that the 'growing pains' of rising social and cultural diversity may be inevitable in the short- to medium-term because diversity tends to challenge social cohesion and lead to greater levels of isolation, prejudice, and/or conflict before its benefits can be reaped in the long-term (Putnam, 2007). However, here it is argued that if rising diversity is managed mindfully, then it can lead to cultural integration and many benefits associated with it sooner rather than later. Chapter 2 will delve more deeply

into the different implications that rising diversity may have for intergroup relations, ranging from cultural separation and marginalisation, to cultural assimilation and integration. Meanwhile, both individuals and societies have a choice to make: They can either choose not to pay attention to issues arising from a more diverse global community, which may exacerbate old conflicts and give rise to new ones; or they can prioritise research and policy related to diversity and its effects, tackle emerging issues associated with diversity, and hopefully reduce social isolation and intergroup tensions. The global community——including scientists, educators, politicians, and business leaders——has an opportunity to reflect on how to effectively manage our increasingly diverse societies (Galinsky et al., 2015).

Different cases can be made in favour of or against allocating more attention to the causes and consequences of diversity. Some have advanced an ethnocentric case, arguing that social homogeneity (i.e., a lack of social diversity) can boost a sense of community and social cohesion, with prime examples being Scandinavian countries such as Denmark and Norway (Eckstein, 1966; Lundvall, 2002). Thus, political parties that are sceptical of immigration have advocated in favour of regulating immigration and growing diversity tightly, rather than supporting and nourishing it (Evans & Mellon, 2016; Franzmann, 2017; Della Posta, 2013; Widfeldt & Brandenburg, 2017). However, this perspective may be too narrow: Because diversity can refer to both cultural (e.g., different ethnic groups) and social (e.g. social class; occupation) categories, it represents a characteristic that is naturally present in any society to a lesser or greater degree. Thus, while the extent of growing diversity could indeed be regulated, the very fact that social diversity naturally exists in most

societies is likely to spawn a myriad of challenges. In order to prevent outcomes such as social or cultural separation or marginalisation, it may be wise to allocate attention and resources to navigating diversity effectively, instead of scapegoating or ostracising certain social or cultural groups. After all, social diversity may not only pose challenges to minorities and/or outgroups (e.g., homeless or poor people, immigrants, refugees), but also majority groups and ingroups themselves (e.g., when conflicts between social groups arise).

What is more, one can make a pragmatic case in favour of nurturing social and cultural diversity. Research suggests that it can add tangible value to societies and organisations, especially if cultural integration is achieved. More precisely, not only does group homogeneity tend to breed narrow-mindedness and groupthink (Janis, 1972; Kerschreiter, Schulz-Hardt, Mojzisch & Frey, 2008; Schulz-Hardt, Frey, Lüthgens & Moscovici, 2000), but a diverse mix of values and ideas can allow people to become better at taking the perspective of others (Hoever et al., 2012; Hoffman, 1959; Sommers, 2006), better innovators, and more creative problem solvers (Crisp & Turner, 2011; Gocłowska, Baas, Crisp & De Dreu, 2014; Simonton, 1997). This so-called 'business case' in favour of diversity is compelling (Herring, 2009), and this thesis contributes new empirical evidence in order to better understand to what extent one might reap cognitive benefits from simple exposure to diversity.

Finally, a moral case in favour of promoting diversity can be made: Conscience and compassion mandate that we attempt to reduce suffering and help people who are different from us, simply because it is the ethically right thing to do (Johns, Green & Powell, 2012; Noon, 2007). Whether or not one is in favour of or against pri-

oritising the challenges associated with diversity on the political and social agenda, many modern societies are becoming multicultural environments that require coalition building to address global challenges. If people are to tackle poverty, intractable conflicts, economic crises, epidemics, and climate change, then they are well advised to deepen their understanding and discourse on how to navigate diversity effectively.

1.3 Aims of the Thesis

The central aim of this thesis is to examine how people adapt to new and complex forms of social diversity. The focus is psychological, in the sense that it will be placed on consequences for affect, motivation, cognition, behaviour. While diversity science has garnered growing research attention in recent years (Crisp & Turner, 2011; Crisp & Meleady, 2012; Mendoza-Denton & España, 2010; Plaut, 2010a), some pressing questions are yet to be addressed. What are the psychological consequences of being exposed to people from varied backgrounds and cultures? How do people psychologically adapt to new forms of diversity over time? With these questions in mind, the empirical work described in this thesis will investigate (1) the idea that exposure to unexpected forms of diversity can influence perceivers' broader cognitive functioning (conceptualised as cognitive reflection in Chapters 4 and 5, and as cognitive flexibility in Chapter 6), (2) the role of individual differences in need for cognition when people learn about new forms of diversity, and (3) how people psychologically adapt to unexpected forms of diversity over time.

1.4 Thesis Overview

Chapter 2 begins by reviewing the current state of the literature on psychological adaptation to diversity. First, psychological adaptation, social and cultural diversity, and counter-stereotypes are defined. Importantly, CSTs are conceptualised as a special type of expectancy violation, which has implications for the theorising in later chapters (particularly Chapter 7). Then, a brief overview of key models and theories is presented and the theoretical progress made in diversity science to date is discussed. The chapter ends with a discussion of empirical and methodological challenges in diversity science, which in turn informs the empirical work conducted as part of this thesis.

Chapter 3 presents a brief overview of the history of the study of counter-stereotypes, which serves as the empirical foundation for the predictions made in later chapters.

Chapter 4 comprises eight experiments testing the hypothesis that exposure to counter-stereotypical diversity can directly impact cognitive reflection.

Chapter 5 presents three experiments testing the hypothesis that individual differences in need for cognition moderate the hypothesised effect of exposure to counter-stereotypes on cognitive reflection.

Chapter 6 presents a preregistered longitudinal experiment examining how people psychologically adapt to exposure to counter-stereotypical diversity over time.

Chapter 7 introduces Expectancy Regulation Theory (ERT), a new theory which

aims to specify how expectancy violations can influence affect, motivation, cognition, and behaviour by triggering one of three possible states: Indifference versus threat (and defensiveness) versus challenge (and open-mindedness).

Finally, Chapter 8 summarises the empirical findings from Chapters 4 to 6 as well as the key insights from Chapter 7, and while keeping the limitations of the presented research in mind, draws theoretical and practical implications for research on psychological adaptation to social and cultural diversity.

Chapter 2

Towards a Diversity Science: Theoretical Progress and Empirical Challenges

---

The purpose of Chapter 2 is to review the current state of the literature on psychological adaptation to social and cultural diversity. I review and discuss the concepts central to this thesis, as well as the main theoretical and methodological approaches in research on psychological adaptation to diversity. Specifically, I first define psychological adaptation, social and cultural diversity, and counter-stereotypes. This is followed by a review of the broad implications of increasing diversity for intergroup relations through the lens of Berry's (1997) model of acculturation. Then, I provide a succinct overview of key models and theories relevant to psychological adaptation to counter-stereotypical diversity, and reflect on the theoretical progress

made to date. The chapter ends with a discussion of empirical and methodological challenges, which in turn informs the empirical work conducted as part of this thesis.

2.1 New Forms of Diversity Challenge Traditional Stereotypes

As a result of social diversity rising, people often come into contact with new and unusual combinations of social, religious, and cultural identities. These days, it is not uncommon to learn about female chief executives spearheading large technology firms, Muslims advocating liberal values such as abortion rights, disabled people winning sports competitions, Catholics coming out as gay, or successful startup founders who did not complete their university studies. Such individuals are exemplars of counter-stereotypical diversity: They do not conform to traditional stereotypes and thus are termed counter-stereotypes ("CSTs"; Crisp & Turner, 2011). People encounter counter-stereotypical individuals when they commute to work, attend school, and pursue leisure activities, or when they travel, change jobs, or move to new places. The media, films, and books frequently feature significant achievements of members of minority groups, such as the film Hidden Figures, which tells the story of three Black female engineers who helped send US American rockets into space. What is more, grassroots social movements have arisen in recent years with the goal to dismantle clichés, for example the #ILookLikeAnEngineer campaign started by a female engineer on Twitter inspiring tens of thousands of people

to share their stories (Anchalee, 2016; Dickey, 2015). Finally, recent history has demonstrated how members of under-represented groups have achieved positions of power and influence, defying expectations and inspiring generations to come, such as Barrack Obama in the 2008 and 2012 US presidential elections ("the Obama Effect"; Marx, Ko, & Friedman, 2009), or Sheryl Sandberg as Chief Operations Officer at Facebook and doyenne of the Lean In movement (Sandberg, 2013). How do people psychologically adapt to new forms of diversity that challenge traditional stereotypes? The following sections define the key concepts in this thesis (psychological adaptation, social and cultural diversity, and CSTs) and present the main theoretical and methodological approaches in research on psychological adaptation to diversity.

2.2 Definitions

2.2.1 Defining "Psychological Adaptation"

Oxford Dictionaries define psychology as "the scientific study of the human mind and its functions, especially those affecting behaviour in a given context" (2018). The field of social psychology in particular has been defined as "an attempt to understand and explain how the thoughts, feelings, and behaviours of individuals are influenced by the actual, imagined, or implied presence of others" (Allport, 1968, p. 3). Importantly, the latter definition emphasises the multilevel nature of psychological processes: They typically have emotional, motivational, cognitive,

behavioural, and/or social causes and consequences. In line with this definition, this thesis will conceptualise experiences as psychological in nature whenever they affect, motivation, cognition, behaviour, and/or social relations.

The term adaptation can imply a process or an outcome (Schmitt & Pilcher, 2004). In the context of biological evolution, adaptation as a process refers to a "creature changing and becoming better suited, or fit, to an environment" (Schmitt & Pilcher, 2004, p. 643). In contrast, adaptation as an outcome can either refer to an attribute that helps a creature survive and reproduce at a particular point in time (Reeve & Sherman, 1993) or to the historical end product of the process of evolution (Williams, 1996). Whereas most evolutionary psychologists examine adaptation as an end product rather than continuous process, this thesis will explicitly focus on psychological adaptation as a *process*, rather than outcome.

Psychological adaptation is therefore presently defined as *discernible changes in emotion, motivation, cognition, or behaviour (or a combination thereof), which people show consistently in response to a stimulus, event, or environment.* For example, an individual who has lived in a small village all their life, but has recently moved to a big city, may be repeatedly exposed to new and unexpected forms of social and cultural diversity: They might meet a Muslim feminist, work with a hippie lawyer, meet a gay married couple raising a child, or make friends with a disabled athlete. If, for example, said individual is politically conservative or has a high need for structure, then they may initially be sceptical of people who challenge stereotypes and social norms, and they may be initially resistant to deviance from common norms (Fay & Frese, 2000; Graham, Haidt & Nosek, 2009; Graham, Nosek

& Haidt, 2012; Inbar, Pizarro & Bloom, 2009; Jost, Federico & Napier, 2009). However, said individual may successfully adapt to new forms of diversity over time as they repeatedly experience contact with outgroup members (Brown & Hewstone, 2005; Christ et al., 2014; Pettigrew, 1998, 2008).

2.2.2 Defining "Social and Cultural Diversity"

Arguably, a comprehensive definition of social and cultural diversity would encompass a plurality of human states (e.g., cognitions, emotions, motivations) and traits (e.g., dispositions, languages). In their article titled "Diversity science: What is it?" Mendoza-Denton and España (2010) advocate in favour of such a broad conceptualisation of diversity by contending that topics that traditionally have been considered as distinct from "diversity" (e.g., relationships, personality, development) often epitomise the diversity between people, groups, and cultures. In line with this idea, the present thesis is based on a broad definition of diversity that encompasses differences in social categories (e.g., gender, ethnicity), personality (e.g., openness to experience, agreeableness), and/or biological characteristics (e.g., sexuality, age).

Whatever the nature of differences between people or groups may be, these differences can either simply co-exist in a pluralistic society, or they can become cross-cutting and complex when combined in new ways. Resonating with existing literature (Crisp & Hewstone, 2007; Crisp & Meleady, 2012; Crisp & Turner, 2011), this thesis therefore conceptualises social and cultural diversity as a phenomenon that manifests itself on a continuum (Figure 2.1).

Figure 2.1: A continuum of social diversity represented by three types of societies, ranging from ancestral monocultural societies, pluralistic societies, to modern socially complex societies.

Specifically, this thesis considers ancestral monocultural societies as located at one end of the diversity continuum. These environments are characterised by a lack of exposure to diversity and represent highly independent communities that are largely self-sufficient (e.g., imagine traditional Himalayan Buddhist villages that are largely disconnected from modern societies). Pluralism is located roughly at the centre of the continuum, representing situations where people from different social and cultural groups co-exist but do not integrate (e.g., Whites, Blacks, and Asians living in segregated neighbourhoods within the same broad area). Most contemporary societies that can be characterised as pluralistic involve a certain level of exposure to diversity, for example when people leave their communities for work or leisure, or when they interact with traditional or digital media. As a result, pluralistic societies that consist of multiple but segregated communities will tend to involve exposure to simple identity structures (e.g., I am White, he is Black), and therefore likely be characterised by relatively low self- and social identity complexity (Linville, 1987; Roccas & Brewer, 2002). Lastly, the other end of the continuum represents complex forms of diversity and refers to environments in which group

affiliations become cross-cutting and inter-twined, and when identities are newly construed (e.g., Muslim feminists, Black female engineers, and gay Catholics living in a racially diverse neighbourhood).

Notably, this dimensional conceptualisation of social and cultural diversity allows us to think about the challenges that diversity poses for psychological adaptation. Ancestral monocultural societies and pluralistic societies represent, arguably, less challenging forms of diversity than socially complex forms of diversity because traditional social roles and categories tend to remain intact, thus echoing the human preference for stability and structure (Caporael, 1997). Complex forms of diversity such as CSTs, on the other hand, may be perceived as relatively more challenging because they conflict with the evolved human propensity for categorical thinking and call into question existing frames of reference (Crisp & Meleady, 2012). Extensive research on multiple and crossed social categorisation supports the plausibility of this argument by showing that forming impressions of surprising category combinations (e.g., an Oxford-educated bricklayer) takes more cognitive effort, more time, and involves more complex reasoning as compared to unsurprising combinations (e.g., an Oxford-educated art critic) (Crisp & Hewstone, 2007; Hutter, Crisp, Humphreys, Waters & Moffitt, 2009). In sum, the experience of diversity can vary in nature and is therefore likely to have divergent implications for psychological processes, depending on the extent of the challenges that the experience poses. In the following section, CSTs and various related constructs will be disambiguated and defined in order to establish conceptual clarity and to set the stage for later discussion.

2.2.3 Counter-stereotypes are a Special Type of Expectancy Violations

As people construe themselves and others in diverse, complex, and crosscutting ways (e.g., "I am both Russian and German, a PhD student, a startup founder), these construals can have wide-ranging implications for psychological, behavioural, and social processes (Crisp & Hewstone, 2007). For example, there is ample evidence to suggest that activating multiple social categories (by asking participants to think of multiple affiliations rather than a single affiliation of an outgroup target) can help reduce intergroup stereotypes and prejudice (Crisp, Hewstone & Rubin, 2001; Hall & Crisp, 2005; Prati, Crisp, Pratto & Rubini, 2016; Prati, Vasiljevic, Crisp & Rubini, 2015b). Because a detailed review of theories and research on multiple social categorisation and its psychological effects is beyond the scope of this thesis, the interested reader is referred to Crisp and Hewstone's (2007) review article delineating when and how multiple social categorisation can lead to more positive intergroup attitudes and more creative, divergent thought. Importantly, the focus of this thesis will be placed on the psychological effects of exposure to counter-stereotypical diversity, that is, information, situations, and events that activate multiple, unusual, and unexpected combinations of social categories (e.g., Oxford-educated bricklayer, female CEO, Muslim feminist, disabled athlete). Put differently, here CSTs are understood as "behaviors and personality traits that buck contemporary cultural norms" (Aubrey & Harrison, 2004, p.116).

In the example of a new city dweller in Section 2.2.1, fixed social categories and

stereotypes would become less functional as the person navigates their environment; instead, considering individuals on a case-by-case basis may be more functional for them (Fiske & Neuberg, 1990). Indeed, one of the key premises in this thesis is that people change the way that they form impressions of each other as they are exposed to new forms of diversity over time: Instead of relying on category-based judgments (e.g., related to gender), they are likely to start making sense of other people based on attributes that go beyond category-membership (e.g., personality traits), or they may combine category membership and individual attributes as sources of information when forming impressions (e.g., by drawing on gender and personality in combination; Fiske & Neuberg, 1990).

However, CSTs represent only one possible way of challenging one's expectations and frame of mind. CSTs can be described as a special type of expectancy violations, which are defined as stimuli or events that contradict people's assumptions about the world (see Chapter 7). Expectancy violations are synonymous with schema violations, which are referred to as "targets and objects that disconfirm schema- and stereotype-based expectancies" (Gocłowska, Baas, Elliot & De Dreu, 2017a, p. 54). One subtype of expectancy or schema violations are social schema violations, which are conceptually superordinate to CSTs. While social schema violations refer to situations when *cognitive schemata about social others* are violated (e.g., an astronaut on the beach, an imam in front of a church), CSTs refer more specifically to situations when *stereotypes* are violated (a gay imam, a Black president). In other words, all CSTs are social schema violations, but not all social schema violations are CSTs. Similarly, the concept of norm violations is related but distinct from CSTs.

While both norm violations and CSTs are examples of expectancy violations, it is notable that the former typically focus on the appropriateness of behaviours in certain situations (Mu, Kitayama, Han & Gelfand, 2015), whereas the latter refer to beliefs about groups that are challenged.

Note that this thesis is explicitly concerned with situations where stereotype-based beliefs are violated, rather than other types of beliefs or expectancies. However, Chapter 7 will discuss the extent to which expectancy violations more generally may have diverging implications for emotion, motivation, cognition, and behaviour. The next section delves more deeply into the implications of increasing diversity for intergroup relations in order to elucidate why it is important to care about this topic. After that, theories of impression formation and psychological adaptation to diversity will be summarised in order to provide the relevant theoretical backdrop for the empirical work conducted in this thesis.

2.3 The Implications of Increasing Diversity for Intergroup Relations

Increases in diversity have implications for both individuals and societies. While exposure to new forms of diversity and intergroup contact can sometimes result in improved intergroup relations (Allport, 1954, 1979) and even cognitive flexibility and creativity (Crisp & Turner, 2011), it may also lead to conflict (Pelled et al., 1999) and acculturative stress (Berry, 1970; Berry, 1997). The consequences of in-

creasing social and cultural diversity are hotly debated in the literature: Broadly speaking, four perspectives have been advanced, reflecting different viewpoints on the psychological and social implications of diversity and intergroup contact. In what follows, the discussion will be organised using Berry's (1997) model of acculturation, where acculturation is defined both as cultural changes in groups and as psychological changes in individuals. Relevant social psychological theories and evidence will be presented and discussed in order to elucidate how rising social and cultural diversity and different types of acculturation strategies may affect intergroup relations. This discussion will provide a broad framework that will help (1) locate the contribution of this thesis and (2) organise the discussion of the empirical findings presented in this thesis in a meaningful way.

2.3.1 Cultural Separation

According to Berry (1970), cultural separation occurs when individuals place value on holding on to their original culture, values, and norms while simultaneously preferring to avoid contact with outgroup members. It can be both a precursor to stress, a consequence thereof, or an active strategy pursued by individuals or groups. Some research suggests that segregation of groups tends to be accompanied by, and may often lead to, prejudice, discrimination, and intergroup conflict (e.g., Campbell, 1965; Esses, Dovidio, Jackson & Armstrong, 2001; Hewstone, Rubin & Willis, 2001; Major, Blodorn & Major Blascovich, 2016; Putnam, 2007). For example, social diversity and segregation may motivate groups to compete for scarce

resources in order achieve or maintain power, prestige, or privilege (Dollard, Miller, Doob, Mowrer & Sears, 1939; Giles & Evans, 1986; Hovland & Sears, 1940; LeVine & Campbell, 1972; Quillian, 1995; Sherif, White, Sherif, Hood & Harvey, 1961), which can be due to many, potentially interacting factors. Amongst others, segregation and intergroup bias can arise as a result of the need for positive self-esteem (Abrams & Hogg, 1988; Tajfel & Turner, 1979), the need to reduce subjective uncertainty (Hogg, 2000, 2007), or the need to preserve oneself and one's culture (Greenberg et al., 1990; Rosenblatt, Greenberg, Solomon, Pyszczynski & Lyon, 1989).

Evidence for the proposition that diversity can be associated with segregation and conflict comes from different settings. Research using experimental games (e.g., the prisoner's dilemma or ultimatum games) has demonstrated that the more that players differ from one another, the more likely they are to defect (or 'cheat'); a finding which has been reported across different countries (e.g., Eckel & Grossman, 2001; Glaeser, Laibson, Scheinkman & Soutter, 2000). Sociological research suggests that greater ethnic heterogeneity can be linked to lower social trust (Anderson & Paskeviciute, 2006; Delhey & Newton, 2005), although a recent comprehensive review of the literature suggests that this effect is highly context-dependent (van der Meer & Tolsma, 2014). Finally, research on workgroups in the United States and Europe suggests that (1) internal heterogeneity (in terms of age, professional background, ethnicity, and other factors) is associated with lower group cohesion, decreased satisfaction, and higher turnover (Jackson et al., 1991; Keller, 2001) and (2) team diversity can be associated with both task and emotional conflict, which in turn have implications for task performance (Pelled et al., 1999). In sum, there

is significant evidence to suggest that social and cultural diversity can be associated with segregation and intergroup conflict.

2.3.2 Cultural Marginalisation

Cultural marginalisation occurs when groups neither have an interest in maintaining their own culture, nor in having relations with other cultures or groups (Berry, 1997), thus breeding social isolation (Putnam, 2007). When cultural marginalisation takes place, it is likely to reduce *social capital*, which is defined as the "links, shared values, and understandings in society that enable individuals and groups to trust each other and to work together" (Keeley, 2007, p. 102). Specifically, Putnam (2007) has theorised that, in the short to medium run, immigration and ethnic diversity may challenge social solidarity and hamper social capital. In other words, "people living in ethnically diverse settings appear to 'hunker down'—that is, to pull in like a turtle. (...) Diversity, at least in the short run, seems to bring out the turtle in all of us" (Putnam, 2007, p. 149-151). However, Putnam also proposes that, in the medium to long run, successful immigrant societies can create new forms of social solidarity that dampen the negative effects of diversity by creating new, more encompassing identities and a broader sense of "we" (an argument that is in line with the common ingroup identity model, see Dovidio, Gaertner, and Saguy (2009)). Taken together, Putnam argues that immigration and social diversity are not merely inevitable, but in the long term also desirable.

What evidence is there that diversity initially results in social marginalisation

and isolation? Putnam (2007) found that diversity makes inhabitants of diverse communities withdraw from collective life and from close friends, distrust their neighbours, have low confidence in their communities and leaders, volunteer less, and withdraw from political activities, even when controlling for potential confounding variables (e.g., economic inequality and crime rates), and ruling out conceptual, statistical, and methodological issues (e.g., alternative explanations, self-selection bias, moderating variables). Other studies similarly suggest that diversity can be negatively related to social cohesion and trust (e.g., Alesina & La Ferrara, 2002). These findings are consistent with classic findings in social psychology on bystander intervention in crises. Specifically, not only have data been presented to suggest that the larger the number of bystanders, the lower the chances that any of them will intervene in an emergency (Darley & Latané, 1968), but researchers have also found that people limit their "span of sympathy" by responding to calls for help along ethnic lines. That is, experimental research suggests that Black participants offered equivalent levels of assistance to both Black and White participants, however White participants helped Black (but not White) participants less frequently (Crosby, Bromley & Saxe, 1980; Gaertner & Bickman, 1971). More broadly, it has been proposed that neighbourhood heterogeneity might make social behaviour less predictable and therefore lead to "information overload", which in turn may result in an inability to process inputs from the environment effectively (Milgram, 1970). To summarise, significant research and theories are consistent with the perspective that diversity can breed isolation, at least in the short term.

2.3.3 Cultural Assimilation

From the acculturation strategies described thus far, it may appear that diversity and acculturation may inevitably lead to social and psychological problems. However, this is not necessarily the case. Cultural assimilation is a process wherein individuals do not wish to maintain their cultural identity but instead desire to assimilate to a new, often superordinate culture or group by adopting their values and norms (Berry, 1997). In response to rising social and cultural diversity, group members can often choose their acculturation strategy——they may or may not decide to assimilate to any given culture. Whereas disciplines such as sociology, economics, and political science have debated and conducted significant empirical research on the consequences of assimilation, only little psychological research has been conducted in this domain (Hornsey & Hogg, 2000). One exception may be seen in research on the common ingroup identity model (Dovidio, Gaertner & Saguy, 2009). The model's key hypothesis bears close resemblance to the idea of cultural assimilation because it suggests that one way to improve intergroup relations is to encourage group members to categorise themselves at the superordinate level (see also Hornsey & Hogg, 2000). For example, someone might be encouraged to think of himself or herself as a European instead of British citizen, thus broadening their collective identity. The goal of this approach is to reduce intergroup bias by systematically altering the perception of group boundaries, which the allows to redefine who is perceived as an ingroup member (Dovidio et al., 2009). The common ingroup identity model explicitly refrains from proposing that superordinate categories should

become over-powering or "all-encompassing". Rather, it is envisaged that in intergroup situations group members retain distinct subgroup identities, but conceive of themselves as "all playing on the same team" (i.e., in terms of having a more inclusive collective identity; Gaertner, Rust, Dovidio, Bachman, and Anastasio 1994, p. 227). In that vein, the common ingroup identity model differs from Berry's (1997) conception of assimilation——the former explicitly encourages distinct subgroup identities, whereas the latter assumes that group members distance themselves from their original social or cultural groups altogether.

Research on the benefits of establishing a common ingroup identity abounds. Laboratory research has found that creating common ingroup identities can help reduce intergroup bias, for example by improving cooperation (Gaertner, Mann, Dovidio, Murrell & Pomare, 1990), increasing positive affect (Dovidio, Gaertner & Loux, 2000), by increasing helping behaviour across group boundaries (Nier et al., 2001), and by stimulating intergroup forgiveness (e.g., that by Jews toward Germans for the Holocaust; Wohl & Branscombe, 2005). Field research in school and professional settings supports the external validity of the model generally finding that stronger perceptions of a common, inclusive identity are associated with improved intergroup relations (Gaertner, Bachman, Dovidio & Banker, 2001; Houlette et al., 2004). However, it is important to keep in mind that creating common ingroup identities (regardless of whether these involve pronounced subgroup identities or not) is not the only way of responding to growing social and cultural diversity. As the next section elaborates, one alternative approach termed cultural integration also holds great promise for improving intergroup relations.

2.3.4 Cultural Integration

People can also respond to diversity by engaging in cultural integration (Berry, 1997). Here, individuals place value on both their original cultures and group memberships, and show interest in new cultures or outgroups. This perspective paints a different picture to some of the perspectives described above: It suggests that exposure to diversity and intergroup contact can be beneficial to individuals and societies (Park & Judd, 2005). This thesis is most closely aligned with this view, assuming that social and cultural diversity can add value to societies if individuals are motivated and able to embrace it.

Perhaps the first theoretical specification echoing the cultural integration perspective was Gordon Allport's (1954) contact hypothesis, which states that contact between members of opposing groups, under the right conditions (such as equal status and common goals), is an effective way to reduce stereotypes and prejudice. In subsequent theoretical work, psychological theorists have both refined and empirically tested new intergroup contact models and theories (Brown & Hewstone, 2005; Dovidio, Gaertner & Kawakami, 2003; Hewstone & Brown, 1986; Pettigrew, 1998). Most notably, Brown and Hewstone (2005) proposed an integrative theory of intergroup contact, which synthesised preceding work. Their theory stresses the importance of the following aspects: (1) In contact situations, group memberships need to be salient if prejudice is to decrease not only towards one outgroup member, but also towards the outgroup as a whole (termed category salience); (2) different dimensions of intergroup contact, such as cross-group friendship (i.e., intimate rela-

tionships between in-group and out-group members), can be differentially powerful depending on the specific context (e.g., the severity of intergroup tensions); (3) different types of mediators, such as empathy, anxiety, and trust, can explain the link between contact and prejudice, and thus they need to be considered carefully when designing targeted and effective interventions.

In light of the significant theoretical progress made, what is the evidence for the proposition that intergroup contact and cultural integration can foster more positive social relations? Meta-analytic evidence suggests that the conditions originally postulated by Allport (1954) are facilitating, but not necessary for the positive effects of contact to occur (Pettigrew & Tropp, 2006; Pettigrew, 2008). It appears that positive contact experiences in their own right are sufficient to reduce intergroup bias (i.e., negative attitudes, emotions, intentions, and behaviours) and to foster social tolerance and solidarity. Large-scale, multilevel studies provide some evidence that intergroup contact is more potent in reducing outgroup prejudice at the neighbourhood level compared with the individual level (Christ et al., 2014). According to these studies, prejudice is not only a function of whom you interact with, but also of where you live. Moreover, research on integration as an acculturation strategy suggests that it (as opposed to separation, marginalisation, and assimilation) incorporates multiple protective factors, such as the willingness for mutual accommodation, involvement in two cultural communities, and the ability to be flexible in personality (Berry, 1997). Finally, new research in the emerging field of diversity science indicates that exposure to counter-stereotypical diversity may promote cognitive flexibility and creativity in domains that are unrelated to

the diversity experience (Di Bella & Crisp, 2016; Gocłowska et al., 2014; Gocłowska & Crisp, 2013; Gocłowska, Crisp & Labuschagne, 2012; Prati et al., 2015b; Prati, Crisp & Rubini, 2015a; Vasiljevic & Crisp, 2013). This thesis aims to conceptually replicate the latter finding in Chapters 4, 5, and 6. Taken together, empirical evidence suggests that social and cultural diversity can be associated with positive psychological and social outcomes, such as trust, tolerance, and creativity. It is notable that cultural integration is typically considered the most adaptive of the possible acculturation strategies in response to social and cultural diversity (Berry, 1997; Berry, 2001).

2.4 Theories of Psychological Adaptation to Diversity

2.4.1 The Continuum Model of Impression Formation

How do people form impressions of others, especially others who they differ from? Fiske and Neuberg's (1990) continuum model of impression formation attempts to answer this question by integrating research on stereotyping with research on person perception. In essence, the model postulates that perceivers can use (1) a target's category membership to form an impression (category-based processes), and (2) specific or unique traits to form an impression (individuating processes). This model suggests that impressions can be formed through both types of processes, and that the distinctions among these processes are matters of degree, rather than

discrete shifts (i.e., it is a continuum, Fiske & Neuberg, 1990). The extent to which a perceiver utilises category-based vs. individuating processes depends on various moderating and mediating variables, which are delineated further below.

At each step of the impression formation process, the perceiver evaluates (or, interprets) whether or not the available information is sufficient to make sense of the target individual. For example, after initially categorising a target (e.g., as Muslim), judging it as relevant to current goals and/or interesting (e.g., he is my colleague), and attending to the target's attributes (e.g., he is a feminist who fights social inequality), the model predicts that the perceiver will attempt to categorise the target in a confirmatory manner (termed confirmatory categorisation and akin to confirmation bias, Nickerson, 1998). However, this process may be unsuccessful (e.g., if the perceiver cannot reconcile how his Muslim colleague can also be a feminist). In this case the perceiver will attempt to recategorise the target (e.g., the perceiver will access a new category: The Muslim colleague has grown up in Western Europe), a process termed recategorisation. But if this process is also unsuccessful because it does not sufficiently help to make sense of the target person, then piecemeal integration is predicted to ensue, which involves an attribute-by-attribute analysis of the target (e.g., the perceiver may ask, "What factors or conditions have motivated my Muslim colleague to become a feminist who fights social inequality?").

The continuum model of impression formation is predicated upon four premises, which specify the nature of impression formation processes, the primary mechanisms underlying them, and the influences of psychological variables (e.g., motivational states or types of information) on impression formation outcomes (Fiske & Neuberg,

1990).

The first premise states that category-based processes typically take priority over individuating (or, attribute-oriented) processes: That is, it is assumed that perceivers primarily use category-based information to form impressions (e.g., ethnicity, age, sex), and that they only use individuated processes (e.g., information related to targets' individual characteristics or circumstances) when necessary. In other words, the extent to which a perceiver progresses along the impression formation continuum depends on the ease with which perceivers can make sense of the target's attributes as fitting an available category (or, an interpretation of fit). If a target's attributes easily fit a category (a pre-existing category or one that the perceiver spontaneously constructs), then category-based impression formation will take place. However, if a perceiver is not able to fit a target's attributes into one or more categories, then the perceiver is more likely to draw on individuating, attribute-oriented processes. Fiske and Neuberg (1990) summarise and discuss research specifying the informational conditions (e.g., expectancy-confirming or -disconfirming information) that tend to elicit different interpretations of the fit between category and attributes.

The second premise specifies the primary mediator through which different impression formation processes are presumed to occur. Fiske and Neuberg (1990) propose that attention is a necessary mechanism, such that increased attention is necessary for more individuating impression formation to take place. In this research, attention is typically operationalised as the amount of time that participants spend considering information about target individuals while making a judgment about them.

The third premise of the model posits that the perceiver's motivation influences the outcomes of impression formation. Specifically, it is assumed that different situations have different interdependence structures——i.e., the target, the perceiver, their interaction, or a third party to the interaction may prompt different motivations (e.g., to form an accurate impression, to reduce feelings of uncertainty). Different motivations, in turn, will determine the likely goal(s) of impression formation and hence push the perceiver either toward the categorising, or the individuating end of the continuum.

Finally, the fourth premise integrates some of the preceding premises, positing that attention and interpretation jointly mediate the effects of motivation and information on how people form impressions. With respect to interpretation, it is important that the information about the target individual is interpreted as diagnostic (i.e., minimally interesting or relevant) for it to trigger categorisation, recategorisation, or piecemeal integration. In addition, Fiske and Neuberg (1990) propose that attention per se does not explain which impression formation route is taken——instead, the nature of the attentional processes is critical because it can be biased or accuracy-driven. If the attentional processes are biased, then the perceiver is more likely to operate at the category-based end of the continuum, but if they are accuracy-driven, then he or she is more likely to operate on the individuating end of the continuum. For example, if a perceiver receives information comprising a category label (e.g., woman) and attributes that are consistent with the category label (e.g., warm, nurturing), then the perceiver is likely to interpret the information as fitting and engage in category-based processing. However, if the perceiver

receives information comprising a category label (e.g., woman) and attributes that are inconsistent with the category label (e.g., dominant, selfish) and if the perceiver is motivated to form an accurate impression of a target (or at least, there is no particular reason to form a biased impression), then individuated processing is likely follow.

The continuum model of impression formation is flexible because it can account for real-life, face-to-face encounters, as well as indirect interactions (e.g., online). It can even be applied to situations that do not involve an interaction at all (for example, when reading about a target individual). The continuum model provides an excellent starting point for understanding the question, "How do people psychologically adapt to social and cultural diversity?" because it distinguishes between two fundamental types of cognitive processing that occur when people learn about social others: Category-based versus individuated processing (otherwise known as system 1 vs. system 2; Crisp & Meleady, (2012); Evans, (2008); Kahneman, (2012); Kahneman, (1973); Strack & Deutsch, (2004)). Finally, the continuum model of impression formation also specifies necessary conditions, mechanisms, and boundary conditions of category-based vs. individuated processing.

2.4.2 The Categorisation Processing Adaptation Generalisation (CPAG) Model

In the early 2010s, two decades after Fiske and Neuberg (1990) published the continuum model of impression formation, research on the psychological consequences

of diversity had started to garner interest again——researchers began considering the broader outcomes of exposure to diversity, aside from the judgements that people make about the person or situation (Crisp & Meleady, 2012; Crisp & Turner, 2011; Mendoza-Denton & España, 2010; Plaut, 2010a, 2010b). Crisp and Turner's (2011) Categorisation Processing Adaptation Generalisation (CPAG) model brought these insights together in an effort to explain how people cognitively adapt to social and cultural diversity. It can be argued that the CPAG model built on Fiske and Neuberg's model by postulating that certain types of diversity experiences can help individuals move away from category-based processing and towards the other end of the continuum, individuated processing.

Crisp and Turner (2011) propose that under a set of conditions, social and cultural diversity is likely to yield beneficial outcomes in a variety of psychological domains that range from intergroup tolerance and self-efficacy to problem-solving and creativity. According to the CPAG model, four conditions must be met for diversity experiences to yield generalised cognitive flexibility (or cognitive transfer effects). First, the model postulates that diversity must be experienced in a way that challenges stereotypical expectations. In other words, the experience must involve exposure to a surprising combination of social categories (e.g., a counter-stereotype, such as an Oxford-educated bricklayer). Second, the individual must be both motivated and able to engage in elaborative processing to resolve the stereotypical inconsistencies (e.g., Why does this person work as a bricklayer if they were educated at Oxford?). Third, the model postulates that these categorisation and processing conditions must be followed by *inconsistency resolution*, a critical

cognitive process that comprises two subcomponents: *Stereotype suppression* and *generative thought*. That is, only if the perceiver discards at least some stereotypes associated with the target (stereotype suppression) and then proceeds to forming individualised impressions of the target (generative thought) will it be possible to adapt to counter-stereotypical diversity. Lastly, the CPAG model suggests that the perceiver will only develop generalised cognitive flexibility if he or she is exposed to stereotype-disconfirming diversity multiple times and repeatedly engages in the inconsistency resolution process just described. It can be argued that the CPAG model extends Fiske and Neuberg's model by postulating that certain types of diversity experiences——i.e., experiences that challenge stereotypes——can help individuals move away from category-based and towards individuated processing.

There is some support for the CPAG model. For example, Gocłowska, Crisp, and Labuschagne (2012) found that thinking of a gender CST (e.g., a female mechanic) boosted creative performance within a short experimental session. In another line of research, Prati, Vasiljevic, et al. (2015b) showed that thinking of CSTs pertaining to gender (e.g., a female mechanic) decreased dehumanisation (i.e., the tendency to consider others as less human than ourselves). Importantly, this change was mediated by a reduced reliance on heuristic thinking, lending support to the model. Finally, research indicates that exposure to CSTs reduces intergroup bias by evoking surprise (Prati et al., 2015a), suggesting that affective-motivational states may play a role in the process of resolving stereotypical inconsistencies.

However, while some predictions of the CPAG model have already been tested, other elements of the model await empirical examination. For example, the model

postulates that perceivers need to be motivated and able to engage in elaborative processing in order to resolve inconsistencies—without these preconditions cognitive adaptation to counter-stereotypical experiences will not ensue. Whereas some initial work has been done on the role of personal need for structure (PNS; which refers to preferences for the desired *outcome* of cognitive activity) in the link between exposure to CSTs and cognitive flexibility (e.g., Gocłowska et al., 2014), it is currently unknown whether need for cognition (i.e., preferences for the desired *amount* of cognitive activity) affects this link. Chapter 5 of this thesis therefore examines the motivation processing condition of the CPAG model by testing whether exposure to CSTs differently affects the cognitive performance of people low versus high in need for cognition. Chapter 6 then presents the first longitudinal test of the CPAG model by repeatedly exposing people to CSTs and measuring their emotional, cognitive, and motivational responses to counter-stereotypical experiences. In the next section, a critical look is taken at interventions capitalising on diversity experiences and special consideration is given to the the question: How can one most effectively design diversity interventions?

2.5 Diversity Interventions

Over the past decades, a tremendous amount of effort has been dedicated to research and interventions aiming to reduce explicit and implicit bias against people from underrepresented or stigmatised groups. For example, the Anti-Prejudice Consortium is an organisation devoted to empowering and educating youth to respect

and embrace differences among all people. Since the consortium's inception in 1997, the Power Over Prejudice (POP) programmes are said to have empowered over 12,000 middle school students and over 1,000 school counsellors in the USA on the importance of embracing diversity in their schools and communities (Anti-Prejudice Consortium, 2016). This section discusses the importance of diversity interventions being evidence-driven, including suggestions for how to design diversity interventions most effectively.

Despite some interventions being evidence-driven, Moss-Racusin et al. (2014) and others (e.g., Devine, Forscher, Austin & Cox, 2012; Paluck & Green, 2009) have observed that most existing diversity interventions and training programs are, in fact, not evidence-based. Worryingly, some interventions may even induce unintended negative consequences such as reactance and backlash, if they imply "that participants are at fault of current diversity challenges" (Moss-Racusin et al., 2014, p. 615; see also Legault, Gutsell, & Inzlicht, 2011). As a result, recent calls have been made in favour of a "scientific approach to the design, assessment, and broad implementation of diversity interventions" (Moss-Racusin et al., 2014, p. 615). Such a scientific approach to diversity interventions——if adopted widely and implemented in social policies——would allow the field of diversity science to have more practical and wide-ranging positive social impact. It would allow scientists and practitioners to reduce intergroup biases and social disparities. In this spirit, the present thesis takes an "interventionist approach" in the sense that it tests a range of possible diversity interventions with the goal to evaluate their effectiveness, reliability, generalisability, and general promise.

Importantly, Moss-Racusin et al. (2014) propose some design elements and measurable outcomes that science-based diversity interventions should incorporate in order to be effective. The authors suggest that the following elements should be components of diversity interventions: Diversity interventions should be (1) grounded in current theory and research, (2) avoid assigning blame or responsibility to participants for current diversity challenges, (3) include a plan for ongoing rigorous evaluation of the intervention's efficacy with different groups, and (4) ideally use active learning techniques so that participants engage with the course content. Additionally, Moss-Racusin et al. highlight the importance of measurable outcomes by stressing that bias literacy should be increased (i.e., participants' awareness of research on diversity issues), participants explicit and implicit biases should be decreased, and that participants' propensity to take action on diversity issues should be increased.

While it may not always be possible to include all of these elements when designing diversity interventions, it is crucial that, at the very least, diversity interventions are evidence-based (i.e., grounded in current theory and empirical evidence). This thesis aims to contribute to an evaluation of the effectiveness and scope of one particular type of diversity intervention: Counter-stereotype (CST) interventions, which will be defined as interventions using CSTs that aim to improve psychological outcomes. In Chapters 4, 5, and 6 empirical evidence will be presented for a range of different CST interventions and their effects on cognitive reflection and flexibility.

2.6 Empirical Challenges to Studying Psychological Adaptation to Diversity

2.6.1 Operationalising Social and Cultural Diversity

Diversity science is an emerging discipline facing both conceptual and methodological challenges. As a result of diversity science being a nascent field, paradigms are needed that allow the study of meaningful diversity experiences in the laboratory, in the field, or in online experiments. In what follows, two paradigms are discussed that can be well suited for the study of diversity. Whereas the first, intergroup contact, is a well-established and prominently used paradigm, the second, exposure to diversity, is a novel paradigm, whose potential is explored in this thesis. The considerations raised in the following sections will inform the hypotheses and approach adopted in this thesis.

2.6.1.1 The Intergroup Contact Paradigm

A wealth of evidence demonstrates at least one route to successful adaptation to diversity: Intergroup contact (Allport, 1954; Brown & Hewstone, 2005; Hewstone & Brown, 1986; Hewstone & Swart, 2011; Pettigrew & Tropp, 2006; Pettigrew, 2008). As described earlier in this chapter, the contact hypothesis states that contact between members of opposing groups, under the right conditions, can be an effective way to diminish stereotypes and prejudice (Allport, 1954). Decades of empirical research on moderators and mediators of the contact hypothesis have been

quantitatively synthesised (Pettigrew, 2008; Pettigrew & Tropp, 2006), leading to the conclusion that the conditions postulated by Allport facilitate positive effects of contact, but are not necessary for them to occur. It appears that the experience of positive contact is powerful enough to reduce intergroup hostilities and promote positive intergroup relations.

Correlational studies examining intergroup contact usually ask participants to self-report the quantity and quality of contact they have experienced with members of different outgroups (e.g., Barlow et al., 2012; Reimer et al., 2017; Schmid, Hewstone & Tausch, 2014; Stathi & Crisp, 2010; Voci & Hewstone, 2003). Notably, while positive contact predicts reduced prejudice, negative contact predicts increased prejudice at a stronger rate (Aberson, 2015; Barlow et al., 2012). However, correlational studies are problematic because they do not allow to infer causal relationships between variables of interest. Therefore, it is important consider experimental evidence on the effects of intergroup contact. In a typical experiment investigating the effects of contact, participants are asked to interact with members of an outgroup over a certain period of time (e.g., Blascovich, Mendes, Hunter, Lickel & Kowai-Bell, 2001; Saguy, Tausch, Dovidio & Pratto, 2009; Sherif et al., 1961; Wilder, 1984; Wilder & Thompson, 1980). In such experiments, the outgroup member in question will often be a confederate who behaves either in a positive or negative manner, which serves to create a positive versus negative contact situation (e.g., Wilder, 1984; Wilder & Thompson, 1980). Researchers have also used naturally occurring intergroup contact to observe what kinds of dynamics can unfold between groups and how intergroup contact influences intergroup attitudes and

behaviours in field experiments (e.g., Sherif et al., 1961).

One interesting advance in contact research is the finding that group members do not actually have to meet face-to-face in order to develop more positive intergroup relations. Indirect forms of intergroup contact, such as extended contact (Wright, Aron, McLaughlin-Volpe & Ropp, 1997) and imagined contact (Turner, Crisp & Lambert, 2007), have been shown effective in reducing intergroup bias. Whereas extended contact refers to the idea that knowledge about an ingroup member having a close relationship with an outgroup member can help decrease intergroup bias, imagined contact is not contingent on the existence of a relationship with an outgroup member. Instead, the mental simulation of a neutral or positive intergroup encounter has been found to promote tolerance and more positive intergroup attitudes (Brambilla, Ravenna & Hewstone, 2012; Crisp & Turner, 2009, 2012; Hayward, Tropp, Hornsey & Barlow, 2017), even in children (Cameron, Rutland, Turner, Holman-Nicolas & Powell, 2011). A recent meta-analysis corroborates the effectiveness of positive and neutral imagined contact in decreasing prejudice (overall $d_+ = 0.35$; Miles & Crisp, 2014), though some boundary conditions have been identified (Hoffarth & Hodson, 2016; Husnu & Crisp, 2010; Visintin, Birtel & Crisp, 2017). Typical mediators of the effect of imagined contact on reduced intergroup bias are decreased intergroup anxiety and increased intergroup trust (Turner, West & Christie, 2013), echoing findings on direct contact.

While there is now a solid evidence base suggesting that different types of contact experiences can nurture more positive intergroup relations, it is important to highlight that diversity interventions can be more or less suited for different con-

texts, depending on their nature. For example, direct contact may not be the best way to help people embrace outgroups and adapt to diversity if (1) communities are highly segregated and there is little opportunity for contact (e.g., the Green Line in Cyprus; Crisp & Turner, 2009); (2) conflict situations are deeply entrenched and thus intractable (e.g., the Israeli-Palestinian conflict; Cohen-Chen, Crisp, Halperin, & Gross 2014); (3) there is little motivation to engage in contact (Crisp & Turner, 2009; Halperin et al., 2012); or (4) any combination of these factors. Ironically, it is precisely where contact interventions may be needed the most that they can backfire, in part because they might exacerbate intergroup bias (Aberson, 2015; Barlow et al., 2012; Hayward et al., 2017), and thus lower the will to engage in meaningful dialogue (Halperin et al., 2012).

For the reasons described above, it is sometimes necessary to take preparatory steps that help pave the way for direct contact at a later stage. In some circumstances it may be helpful to develop interventions to promote the belief that groups can change, which can boost hope and the willingness to compromise for peace (Cohen-Chen et al., 2014; Halperin et al., 2012). Alternatively, it may be useful to develop interventions that induce paradoxical thinking (i.e., by presenting new but extreme information that makes the individual perceive their current beliefs or situation as irrational or senseless) before encouraging direct contact, which can have various advantages compared to both contact and counter-stereotype interventions (Hameiri, Porat, Bar-Tal, Bieler & Halperin, 2014). Lastly, indirect interventions, such as imagined contact, show particular promise whenever direct contact might be experienced as overwhelming and people need "mental space" to prepare for an

intergroup encounter in a psychologically safe way. The imagined contact paradigm can allow individuals to get accustomed to the idea of meeting an outgroup member without the risks involved in direct contact. This thesis capitalises on this argument and utilises the imagined contact paradigm to test whether imagining contact with counter-stereotypical individuals can have a broad impact on cognitive functioning (see Chapter 4), in addition to the well-established finding that imagined contact can help reduce intergroup bias (Miles & Crisp, 2014).

2.6.1.2 The Exposure Paradigm

The contact paradigm, regardless of whether contact is imagined or direct, is powerful and has many advantages; however, it has some further drawbacks besides the ones discussed above. The contact paradigm makes it relatively easy to study pluralism (e.g., A meets B), but it can be rather difficult to study complex, counter-stereotypical diversity (e.g., A meets BKTY, see Section 2.2.2). This is because CSTs are, by definition, unusual individuals who tend to challenge norms and deviate from the status quo. Social outliers, in turn, tend to be evaluated more negatively compared to those who conform to prevalent norms (called the "deviant bias"; Rubin, Paolini, & Crisp, 2010; Rubin, Paolini, & Crisp, 2011; Rubin, Paolini, & Crisp, 2013). Thus, asking participants to meet (or imagine meeting) an expectancy-violating individual (for example, a deviant or misfit) may be experienced as threatening, intimidating, or overwhelming (Förster, Higgins & Werth, 2004; Mendes, Blascovich, Hunter, Lickel & Jost, 2007).

With the potential negative consequences of contact (direct or indirect) in mind,

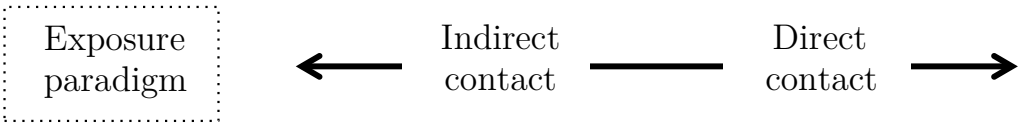

Figure 2.2: The contact continuum (on the right) supplemented by the exposure paradigm (on the left) as a preparatory, pre-contact intervention.

the present thesis also investigates the impact of a preparatory step that involves learning information about an outgroup member or counter-stereotypical individual, but does not require imagining meeting or actually meeting somebody—I term this an 'exposure paradigm' (see Figure 2.2). Introducing an exposure paradigm allows scientists to investigate contexts that are characterised by high threat (Mendes et al., 2007) or by high psychological distance (Trope & Liberman, 2010), such as media or online environments, and it potentially allows to develop scalable online and offline interventions that can help people psychologically adapt to social and cultural diversity (cf. Paunesku et al., 2015). In sum, the present thesis uses both the intergroup contact and exposure to diversity paradigms to test key predictions of the CPAG model.

2.6.2 Conceptualising and Operationalising Cognitive Functioning

Section 1.3 introduced the idea that diversity experiences can impact people's broader cognitive functioning. However, what exactly does the term 'cognitive func-

tioning' encompass? There are myriad ways in which one could conceptualise or operationalise cognitive functioning, ranging from concepts such as executive functioning (i.e., higher order cognitive operations that are involved in the planning, execution, and regulation of behaviour; Baddeley, 1996; Baddeley & Hitch, 1974; Norman & Shallice, 1986), cognitive reflection (Frederick, 2005), learning and cognitive development (Ericsson, Krampe & Tesch-Römer, 1993; Fischer, 1980; Piaget, 1964), dimensions of intelligence or cognitive ability (e.g., verbal comprehension, quantitative reasoning, visual-spatial reasoning; Ritchie, 2015; Terman & Merrill, 1972; Wechsler, 1939), creativity (i.e., the production of ideas or outcomes that are both novel and appropriate to some goal; Amabile, 1983; Simonton, 2000), to cognitive or integrative complexity ("the capacity and willingness to acknowledge the legitimacy of competing perspectives on the same issue (differentiation) and to forge conceptual links among these perspectives (integration)"; Tadmor, Tetlock, & Peng, 2009, p. 105; see also Benet-Martínez, Lee, & Leu, 2006 and Maddux et al., 2014). The following section delves more deeply into two classes of processes that represent different types of cognitive functioning: (1) executive functioning and (2) cognitive flexibility and creativity. Although these types of cognitive processes have been previously investigated as outcome variables in research on social and cultural diversity, this thesis builds on this work by exploring boundary conditions, mechanisms, and longitudinal effects of exposure to CSTs on cognitive functioning (specifically, cognitive reflection and flexibility).

2.6.2.1 Executive Functioning

Executive functioning, broadly speaking, helps people to plan, execute, and regulate their behaviour (Baddeley, 1996; Baddeley & Hitch, 1974; Macrae, Bodenhausen, Schloerscheidt & Milne, 1999; Norman & Shallice, 1986). Norman and Shallice (1986), as well as Baddeley (1996), proposed that existing cognitive schemata operate automatically based on habits in order to control action. However, when expectancies are violated or when individuals are confronted with something novel, a supervisory attentional system is assumed to step in and take on executive control. This characterisation of processes is comparable to dual-process accounts of reasoning, which distinguish between cognitive processes that are fast, automatic, and unconscious versus slow, deliberative, and conscious (also called heuristic vs. systematic processing; Crisp & Meleady, 2012; Evans, 2008; Kahneman, 2012; Kahneman & Tversky, 1973; Stanovich & West, 2002; Strack & Deutsch, 2004). Executive functioning arguably encompasses the construct of cognitive reflection, which is defined as "the ability or disposition to resist reporting the response that first comes to mind" (Frederick, 2005, p. 36). The ability to resist acting upon an intuition requires individuals to engage top-down regulatory processes. Consistent with this idea, the process of inconsistency resolution that can be triggered when somebody encounters a counter-stereotypical individual has itself been characterised as an executive function because it is assumed to help suppress existing schemata and to help generate new impressions of expectancy-violating individuals, stimuli, or events (Crisp & Hewstone, 2007; Crisp & Turner, 2011; Macrae et al., 1999). In the present

work, cognitive functioning is mostly——albeit not exclusively——operationalised as cognitive reflection, which is measured via the Cognitive Reflection Test (CRT; Frederick, 2005; Toplak, West, & Stanovich, 2014; see Chapters 3 and 4).

2.6.2.2 Cognitive Flexibility and Creativity

Exposure to counter-stereotypical diversity not only requires the engagement of executive functions to make sense of expectancy-violating stimuli or events, but more recent models (in particular, the CPAG model) suggest that exposure to CSTs can even boost cognitive functioning in domains beyond the expectancy violation itself. Specifically, recall that, according to the CPAG model, exposure to CSTs, under the right conditions, can enhance domain-general cognitive flexibility. Cognitive flexibility is typically defined as the "(...) capacity to 'break set', go beyond the established and mentally accessible ways of thinking in favor of thinking differently from other people or differently from what is habitual" (Gocłowska & Crisp, 2013, p. 218). Empirical evidence for the prediction made by the CPAG model has been presented, both with respect to cognitive reflection, and with respect to cognitive flexibility (e.g., Gocłowska et al., 2014; Gocłowska, Crisp, & Labuschagne, 2012; Prati, Vasiljevic, et al., 2015b; Vasiljevic & Crisp, 2013). However, as this research is still in its early stages, a more systematic investigation of possible mechanisms, moderators, boundary conditions, and longitudinal effects of exposure to CSTs on cognitive reflection and creativity is needed. The present thesis aims to contribute to this systematic investigation. In Chapter 6, cognitive functioning is conceptualised as cognitive flexibility and operationalised as an idea generation task that requires

participants to come up with new names for a product or item (cf. Gocłowska et al., 2014).

2.7 Summary

Chapter 2 has defined the key constructs in this thesis——psychological adaptation, social and cultural diversity, counter-stereotypes, and cognitive functioning——and introduced models and theories that may help to understand how people psychologically adapt to counter-stereotypical experiences. The continuum model of impression formation (Fiske & Neuberg, 1990) helped to specify the necessary conditions, mechanisms, and boundary conditions of category-based versus individuated processing in the context of person perception, thus providing the conceptual foundation of this thesis. The CPAG model (Crisp & Turner, 2011) provided some of the key hypotheses that this thesis aims to test by deriving concrete predictions regarding how people adapt to social and cultural diversity. Specifically, this thesis is going to test (1) whether exposure to counter-stereotypical diversity can directly change cognitive reflection/flexibility (Chapters 4, 5, and 6), (2) the role of individual differences in need for cognition in the effect of exposure to CSTs on cognitive reflection (Chapter 5), and (3) the effects of repeated exposure to CSTs on cognitive flexibility and intergroup bias (Chapter 6). Whereas Chapter 4 employs interventions based on imagined contact with counter-stereotypical individuals, Chapters 5 and 6 draw on the exposure paradigm, simply presenting participants with information about counter-stereotypical individuals. In the next chapter, a brief history of

the study of CSTs is presented and discussed, which forms the empirical foundation of this thesis.

Chapter 3

A Brief History of the Study of Counter-stereotypes

The aim of this chapter is to provide a brief overview of the history of the study of CSTs from a social psychological perspective. I start by reviewing work on the effects of counter-stereotype interventions on explicit and implicit bias, and then proceed to work investigating the effects on information recall and impression formation. Throughout, I reflect on the strengths and weaknesses of different studies in order to evaluate their evidential value. This work represents the empirical foundation of the chapters to follow.

3.1 Introduction

In an article titled "The Quiet German" from December 2014, George Packer from The New Yorker magazine chronicled the rise of German Chancellor Angela Merkel, describing her as the most powerful woman in the world (Packer, 2014). In this piece, one paragraph in particular stands out, highlighting how atypical Merkel is in the context of German politics:

> "Among German leaders, Merkel is a triple anomaly: a woman (divorced, remarried, no children), a scientist (quantum chemistry), and an Ossi (a product of East Germany). These qualities, though making her an outsider in German politics, also helped to propel her extraordinary rise."

The extent to which these unusual characteristics helped versus hindered Merkel's career is debatable. Notwithstanding, it is important to realise that she was the first woman in Germany to attain the position of Chancellor, which she has been able to maintain for the past decade. Clearly, she represents a counter-stereotype, in a similar way to Barack Obama (see section 2.1). Merkel and Obama are counter-stereotypical individuals because they defy and disconfirm stereotypes that are typically associated with groups they belong to (i.e., women and Blacks). In what follows, a brief history of the study of CSTs is presented, which will provide the relevant theoretical backdrop for the experiments introduced thereafter.

3.1.1 Counter-stereotypes and Intergroup Bias

3.1.1.1 Effects on explicit bias

Some of the earliest work on exposure to CSTs can be traced back to research on intergroup contact and its effects on explicit bias (i.e., outgroup evaluations) and stereotype change. For instance, Wilder (1984) examined the extent to which the perceived typicality of an outgroup member would be effective in improving attitudes and changing stereotypes related to their group. In one experiment, Wilder created an intergroup scenario by asking students of rival colleges to interact with a confederate who was (allegedly) either part of the ingroup versus the outgroup college, pleasant versus unpleasant, and typical versus atypical of the outgroup college. The author found that compared with other conditions, contact with a pleasant and typical outgroup member helped improve evaluations of the respective outgroup the most. However, a second experiment revealed an important caveat: A pleasant interaction with a typical outgroup member did not improve outgroup attitudes when the outgroup member actively displayed a negative stereotype directed towards the ingroup. In other words, the negative stereotype seemed to undermine the positive, pleasant interaction and prevent outgroup attitudes from improving. A third and final experiment examined why typical and pleasant contact had a stronger impact on outgroup evaluations than atypical and pleasant contact. Wilder discovered that successful contact with a typical and pleasant outgroup member generalised to evaluations of the outgroup as a whole because the typical outgroup member was viewed as more predictive of other outgroup members. Remarkably, while contact

with an outgroup member influenced attitudes towards the outgroup as a whole in all three experiments, it had virtually no effect on stereotypes of the outgroup. Various accounts could explain this finding: Either a single contact experience was insufficient to change stereotypes and the predicted effects may only be observed after repeated contact experiences; alternatively, the individual outgroup member may have been considered an exception and thus "subtyped"; lastly, contact with the outgroup member may have reduced the likelihood of stereotypes being applied in that particular setting rather than changing stereotypes in general. Further, aside from these theoretical arguments, an important methodological limitation needs to be raised: The experiment in Wilder (1984) tested only female participants and had very small sample sizes (approx. 10-12 participants per cell). This significantly limits the external validity of the experiments and suggests that they may have been underpowered, thus likely failing to detect effects that are small or medium-sized. In addition, it would have been useful to know the effect sizes for the reported effects in order to be able to evaluate the practical significance of the findings. In conjunction, these early experiments by Wilder do, however, seem to suggest that contact between an ingroup and outgroup member can improve attitudes towards the outgroup as a whole as long as the interaction is positive, the outgroup member is typical, and no negative stereotypes about the ingroup are implicated.

Subsequent research has allowed to refine Wilder's conclusion by showing that sometimes exposure to, or contact with, atypical individuals can, in fact, reduce intergroup bias (e.g., Kunda & Oleson, 1995; Rothbart & John, 1985; Wilder, Simon, & Faith, 1996). The work by Wilder et al. (1996) in particular suggests that counter-

stereotypical information can change explicit stereotypes about outgroups if (1) the disconfirmer is otherwise perceived to be a typical outgroup member, and (2) the attributions made to account for the outgroup member's counter-stereotypical behaviour are dispositional (rather than motivational or external) in nature. Taken together, this research provides some evidence that exposure to CSTs can reduce explicit intergroup bias and in some circumstances even result in stereotype change (for a review see Hewstone, 1994). In Chapter 6, the hypothesis that exposure to CSTs can reduce explicit intergroup bias (i.e., explicit prejudice) was tested in a longitudinal experiment spanning two weeks.

3.1.1.2 Effects on Implicit Bias

Other research has examined whether forming counter-stereotype intentions can help reduce implicit bias (i.e., the automatic activation of stereotypes and prejudice; Blair & Banaji, 1996). Here, counter-stereotype intentions were operationalised via a semantic priming task: Participants completed a series of trials and were instructed to expect stereotypical vs. counter-stereotypical targets following certain regularly occurring words. Participants in the stereotypical condition received a 5:3 ratio of stereotypical to counter-stereotypical trials, and vice versa for participants in the counter-stereotypical condition. Hence, forming a stereotype vs. counter-stereotype strategy was designed to help participants improve their performance. Blair and Banaji observed that forming a counter-stereotype strategy significantly reduced the automatic activation of gender stereotypes and under some circumstances even led to a complete reversal of stereotypical responses. This finding was conceptually

replicated in subsequent research investigating whether counter-stereotypical mental imagery can reduce implicit bias (Blair, Ma & Lenton, 2001; Kawakami, Dovidio, Moll, Hermsen & Russin, 2000). Across five experiments, Blair, Ma, and Lenton (2001) found that participants who engaged in counter-stereotypical mental imagery (e.g., imagining a strong woman) showed significantly weaker implicit stereotypes than participants who engaged in neutral, stereotypical, or no mental imagery. In this research, implicit stereotypes were measured with the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) as well as by means of two measures based on signal detection theory and the false memory effect. Similarly, Dasgupta and Asgari (2004) found in two studies——a laboratory experiment and a year-long field study——that exposure to counter-stereotypical women leaders resulted in women expressing fewer automatic stereotypical beliefs about their ingroup; this effect was mediated by the frequency of exposure to women leaders. The researchers concluded, "(...) these findings underscore the power of local environments in shaping women's nonconscious beliefs about their ingroup" (Dasgupta & Asgari, 2004, p. 642). Recent research on CSTs and intergroup bias more generally supports the effectiveness of CST interventions in reducing implicit bias. In a comparative investigation of 17 interventions designed to reduce implicit racial preferences (e.g., appeals to egalitarian values, perspective taking, imagined contact, imagining vivid counter-stereotypical scenarios), Lai et al. (2014) discovered that only eight of the 17 interventions were effective at reducing implicit preferences for Whites compared with Blacks (total $N = 17{,}021$). Notably, the most potent intervention turned out to be vividly imagining a counter-stereotypical scenario, with an average meta-analytic

effect size of $d_+ = 0.49$. In sum, research on CSTs and implicit bias lends support to the idea that stereotype activation can be malleable (rather than intractable) and that intentional processes, such as counter-stereotypical mental imagery, can reduce automatic stereotype activation.

3.1.2 Counter-stereotypes and Impression Formation

In parallel to the work on CSTs and intergroup bias described above, the 1980s and 1990s saw a surge in research interest with respect to the effects of CSTs on impression formation. When people form impressions of other people, then on the one hand, they need to encode and organise the social information in memory (Srull & Wyer, 1989). On the other hand, they often also need to transform the resulting mental representation (i.e., impression) into social judgements and behavioural decisions (Srull & Wyer, 1989). An early study by Hastie and Kumar (1979) examined the effects of congruent versus incongruent information on person memory. A sample of undergraduate students received information about another person in the form of trait adjectives and sentences describing their behaviour, and was then asked to form an impression about the person as well as to recall the sentences. The results revealed that recall was higher for behaviours that were incongruent (vs. congruent) with a personality trait. However, subsequent research qualified this finding by showing that incongruent (vs. congruent) items were only more likely to be recalled when they could be attributed to dispositional rather than situational causes (Crocker, Hannah & Weber, 1983). Additionally, the research by Crocker

and colleagues revealed that people prefer to make situational attributions when information is inconsistent, and dispositional attributions when information is consistent. This suggests that Hastie and Kumar's (1979) findings may be limited to situations in which situational attributions for inconsistent information are not accessible. Thus, it appears that surprising, inconsistent information does not per se carry a recall advantage; rather the recall advantage hinges on certain conditions. This early research has provided initial insight into how people remember conflicting social information.

3.1.2.1 Emergent Attributes and Generative Thought

Subsequently, researchers have not only attempted to understand how perceivers recall social information, but also how they form social judgements. For example, in two studies, Hastie, Schroeder, and Weber (1990) asked participants to list attributes that characterise simple versus incongruent social category labels (e.g., mechanic vs. female mechanic). In Study 1, the authors aimed to capture a 'snapshot' of the contents that participants would produce when learning about stereotypical versus counter-stereotypical individuals. A total of 140 US American undergraduate participants were presented with three sets of four social category labels, which were paired to create stereotypical versus counter-stereotypical exemplars. For example, the labels 'Black', 'Jewish', 'janitor', and 'banker' were combined to yield different category conjunctions, such as 'Black janitor' and 'Jewish banker' (common, stereotypical), and 'Black banker' and 'Jewish janitor' (rare, counter-stereotypical). Hastie and colleagues discovered that participants who re-

sponded to surprising category conjunctions produced "emergent attributes" that had not been mentioned previously, and were thus less confined by existing category labels. In some cases, over 40% of the generated attributes had not been mentioned for either of the constituent categories, and where thus emergent. For example, participants generated 42% emergent attributes on average for the conjunction 'male nurse' and 49% of emergent attributes for the conjunction 'Republican social worker', but only only 8% emergent attributes for 'feminist social worker' and 18% emergent attributes for 'male mechanic'. To conceptually replicate these findings, Hastie, Schroeder, and Weber (1990) conducted a second study, which used the same materials as Study 1, but instead of asking participants to list attributes, they were instructed to rate the target individuals on bipolar trait-adjective scales (e.g., ambitious/unambitious, warm/cold, intelligent/unintelligent). Nighty-eight US American undergraduate participants were recruited and asked to indicate their ratings on the bipolar scales for both simple categories (e.g., 'female', 'mechanic'), and category conjunctions (e.g., 'female mechanic'). After that, participants were asked to reflect on their reasons for giving certain ratings. The researchers interpreted the scale responses by calculating the bounds as defined by the two ingredient ratings. That is, a participant's bounds would be defined by the occupation rating (e.g., banker or mechanic) and the background category rating (e.g., gender or ethnicity). All ratings that fell outside these bounds were considered 'outside ratings' and interpreted as corresponding to emergent attributes because participants had gone "(...) 'out of the bounds' of values defined by the two ingredient categories" (Hastie, Schroeder & Weber, 1990, p. 244). The results revealed that, on average,

28% of the ratings were outside the range defined by the occupation and background ratings, and that this percentage was particularly high for incongruent conjunctions. When the authors examined participants' explanations for their ratings in order to find clues regarding the psychological processes involved, they uncovered that 10% of responses entailed "complex, deeper forms of reasoning about the conjunction" (p. 245). For example, participants sometimes reflected on conditions necessary to enter an occupation (e.g., 'As a woman, you've got to be dominant in order to survive'), sometimes on the effects of being an unusual member in a occupation (e.g., 'As a woman, you've got a lot more to prove than a man in the same job'), and sometimes referenced subtypes (e.g., 'Male nurses probably need to fill a special role, for example when their physical strength is required'). Taken together, Hastie and colleagues uncovered one possible overarching psychological process in response to counter-stereotypical exemplars: Learning about incongruent categories appeared to make people engage in a complex, creative reasoning process (termed generative thought in the CPAG model, Section 2.4.2) that served to explain why or how the unexpected information made sense.

The two studies just described are rigorous in the sense that they had relatively high sample sizes and well-controlled study materials. The latter, in particular, allowed to hold the content pertaining to the target individuals constant by crossing categories. Limitations of these studies include their within-subjects nature, the rather biased participant samples, and the subjective nature of the content-coding scheme. Regarding the within-subjects design, participants rated both stereotypical and counter-stereotypical category conjunctions in the same session, which may

have anchored or biased their judgements (e.g., seeing congruent categories first may have led to incongruent categories being perceived as more extreme). In order to test more precisely how people respond to CSTs, one would preferably randomly assign participants to a stereotypical versus counter-stereotypical condition and then analyse how their judgements contrast with each other. The fact that the participant samples were US American undergraduate students limits the generalisability of the findings to other populations and cultures. Also, it is unclear whether only one or multiple coders content-analysed participants' reflections. It is important that future research employs multiple coders in order to determine the extent of inter-rater reliability and to minimise the influence of subjective biases in content-coding, and that it replicates the findings with more representative samples in other contexts and cultures.

In a related line of research, Kunda, Miller, and Claire (1990) similarly found that participants who were exposed to social categories with conflicting implications (e.g., Harvard-educated carpenter) produced more emergent attributes and showed a higher level of causal reasoning as compared to participants who were exposed to simple social categories (e.g., carpenter). More precisely, in four studies the researchers examined whether combinations of social categories, especially those that have conflicting implications, would lead to causal reasoning and the desire to generate explanatory hypotheses. In Study 1, the authors aimed to demonstrate that people tend to form unified images of persons who belong to surprising combinations of social categories by drawing on causal reasoning. They conducted the study with 85 undergraduate participants at a US American college in a classroom setting,

asking them to describe different target persons. Participants were split into three groups: (1) those describing counter-stereotypical target individuals (e.g., Harvard-educated carpenter), (2) those describing one of the categories (e.g., someone who is Harvard-educated), and (3) those describing the remaining category (e.g., someone who works as a carpenter). Participants were also asked to rate how surprised they would be to hear about a person belonging to either of the categories or the conjunction of categories. Kunda et al. (1990) discovered that most participants were able to form an impression of counter-stereotypical individuals fairly easily, and importantly, that participants in the counter-stereotypical group generated a variety of causal narratives and emergent attributes in order to make sense of unusual target individuals. For example, participants described a Harvard-educated carpenter as non-conformist and non-materialistic. The researchers also correlated average surprise ratings with the number of causal antecedents (i.e., descriptions referring to causes of group memberships) and found a moderately high correlation ($r = .46$). They concluded from this finding that surprise may have triggered causal reasoning, although it should be noted that this evidence is correlational in nature, so causality should not be assumed. These initial data reported by Kunda and colleagues therefore tentatively suggest that people may rely on causal explanations in order to form unified impressions of counter-stereotypical individuals. While these are intriguing findings, the conclusions that can be drawn from them are limited because they were primarily based on qualitative research methods——some summary statistics were provided, but inferential statistics were not used to compare the different groups of participants in terms of their descriptions, emergent attrib-

utes, and/or ratings. For example, it would have been useful to directly compare the number of emergent attributes generated *across groups* to see if this difference was statistically or practically significant. Next, in Study 2, Kunda and colleagues explored how exactly people try to resolve conflicting social categories. For example, would they resolve them through averaging (e.g., by asking "How likely is intelligence to be for a Harvard-educated person, and how likely is it for a carpenter?", and then averaging the two estimates?). Or would one constituent typically dominate so that the emergent attribute would inherit the attributes of this dominant constituent? The authors hypothesised that neither strategy would be particularly pervasive, bur rather that a mix of these strategies would be employed. They recruited 98 participants from the same college as in Study 1 and using the descriptions generated previously, they asked participants to rate two counter-stereotypical combinations (Harvard-educated carpenter and blind lawyer) regarding how likely each target would be to possess a range of emergent attributes (e.g., non-conformist, non-materialistic). The results showed that the ratings of counter-stereotypical individuals were not viewed as descriptive of members of the constituent categories. From this, Kunda and colleagues concluded that information *external to the constituent categories* was used to form impressions of combinations. Further, they found that participants indeed used a mix of strategies to make sense of the CSTs: Each category conjunction contained at least three attributes that were taken from each constituent category, and at least three attributes were averaged across the constituents. Not once did a constituent completely dominate the combination, nor was averaging the prevalent strategy. In sum, the results suggest that the inconsist-

ency resolution strategies were relatively diverse and dependent on each category combination. To corroborate these conclusions, Studies 3 and 4 conceptually replicated the findings of Studies 1 and 2 with larger undergraduate samples from a different university. Taken together, Kunda, Miller, and Claire (1990) concluded that "there is extensive evidence for the role of causal reasoning in the combination of social categories" (p. 572). Notably, however, the researchers neither tested this assertion in a more representative sample of the US population, nor outside the US American context. Thus, these limitations resemble the ones discussed earlier pertaining to Hastie, Schroeder, and Weber (1990).

The work by Kunda, Miller, and Claire (1990) differs from Hastie, Schroeder, and Weber (1990) in its emphasis of which psychological processes are critical: The former emphasise the surprise and puzzlement aroused by exposure to CSTs and the desire to resolve the puzzlement through causal reasoning that draws on broader world knowledge, whereas the latter describe a broader array of possible psychological processes in response to CSTs (i.e., participants reasoning about the possible causes, consequences, and implications of the being a counter-stereotypical individual). Whereas the studies reported by Hastie and colleagues primarily utilised a within-subjects design, the studies reported by Kunda and colleagues involved a between-subjects design that allowed to compare the responses of different participant groups (which the authors only occasionally capitalised on). In conjuction, these two lines of research, despite their limitations, lend reasonably strong support for the hypothesis that exposure to CSTs involves an impression formation process that typically draws on causal and creative reasoning (i.e., generative thought).

3.1.2.2 The Suppression of Constituent Attributes

Building on the above insights, subsequent research has uncovered that, in addition to generating emergent attributes, inconsistency resolution also encompasses a second component process——it requires people to suppress constituent attributes (Hutter & Crisp, 2005). Recall that in typical CST impression formation paradigms, constituent attributes are attributes that participants produce for both a surprising category combination as well as for its constituent categories. For example, the attribute "hard-working" may describe both a Harvard-educated carpenter, as well as someone who is Harvard-educated, or someone who is a carpenter; it thus represents a constituent attribute. In contrast, the attribute "non-conformist" may describe a Harvard-educated carpenter, but it is unlikely to apply to someone who is either Harvard-educated or a carpenter; it thus represents an emergent attribute. Based on three experiments, Hutter and Crisp (2005) concluded that inconsistency resolution involves both the inhibition of already activated attributes *and* the creative generation of new attributes. Specifically, Hutter and Crisp first established in a pilot test involving 36 participants that the category conjunctions to be employed in the experiments (Oxford-educated art critic vs. bricklayer) were perceived as surprising and unfamiliar. Then, in Experiment 1 they recruited 38 undergraduate participants, who were randomly allocated to one of four conditions: Generating constituent versus emergent attributes for a familiar (Oxford-educated art critic) versus unfamiliar category conjunction (Oxford-educated bricklayer), which were embedded in excerpts ostensibly taken from a job application. Participants were

instructed to spontaneously list up to 20 attributes describing the category combinations within 120 seconds. In addition, a further 60 participants were asked to generate attributes that came to mind when thinking about either of the three constituent category labels (i.e., Oxford-educated, art critic, bricklayer). This served to provide a baseline against which it was compared whether the attributes listed for the category conjunctions were inherited from the constituents versus emergent. Following a procedure described in Hastie, Weber, and Schroeder (1990), Hutter and Crisp first had two independent coders carry out a redundancy check, which helped ensure that similar concepts were only counted once (e.g., "artistic" and "arty" were both coded as "artistic"). After that, the two coders classified each attribute listed by the participants as either constituent or emergent based on the definitions described above; interrater agreement was acceptable and thus the coders' ratings were averaged. The results showed that significantly fewer constituent and significantly more emergent attributes were used to define the unfamiliar (vs. familiar) category conjunction, thus rendering initial support for the hypothesis that suppressing constituent attributes may play an important role when people form impressions of counter-stereotypical individuals. In Experiment 2, Hutter and Crisp (2005) aimed to conceptually replicate these findings by simply presenting the category conjunctions without embedding them in additional text. Otherwise, the method in Experiment 2 was identical to Experiment 1. The authors recruited 120 undergraduate participants who were randomly assigned to the same four conditions as before. In short, the authors replicated the effects of Experiment 1——when asked to form an impression of an unfamiliar (vs. familiar) category conjunction,

participants generated fewer constituent and more emergent attributes. In the final and third experiment, Hutter and Crisp established the generality of the effect with new categories. After having confirmed that the new counter-stereotypical category conjunctions were again perceived as more surprising and less familiar in a pilot study, the authors recruited 80 undergraduate participants who were randomly assigned to a 2 (target gender: male vs. female) x 2 (occupation: nurse vs. mechanic) x 2 (attribute: constituent vs. emergent) design. The remaining procedure was identical to the previous experiments. The analyses revealed that the predicted effects were replicated for female, but not male target individuals. The authors speculated that this may have been due to the fact that their participants on average perceived male nurses as much less surprising than female mechanics. In conjuction, the experiments presented by Hutter and Crisp (2005) can be said to provide significant evidential value regarding the role of suppressing constituent attributes as part the inconsistency resolution process: The experiments were well-designed, reasonably well-powered (assuming a large effect), and they were demonstrated in English-speaking samples in the United Kingdom. However, the experiments could be improved if they were replicated in a sample more representative of the general population, in different cultures, and with additional unfamiliar category conjunctions. In addition, it would have been useful to know the effect sizes for the reported effects in order to evaluate the practical significance of the findings. To summarise, given the evidence presented it does appear that inconsistency resolution prompts people to suppress constituent attributes, in addition to generating emergent attributes.

3.2 Summary

The research reviewed in this chapter shows how exposure to CSTs can affect explicit bias, implicit bias, and impression formation (i.e., information recall and inconsistency resolution). It lays the foundation for the work presented in this thesis as it suggests that counter-stereotypical information can help reduce intergroup bias, that it often commands attention, and that it can spur deeper and more creative cognitive processing related to the CST individual. However, more recent models and research suggest that the effects of CSTs can extend beyond the intergroup domain and the CST individual, in line with predictions by the CPAG model. In the following chapter some early research that has tested the CPAG model is briefly summarised, which in turn provides the context for eight new experiments presented in Chapter 4.

Chapter 4

Imagined Counter-stereotypical Contact and Cognitive Reflection

In this chapter I present eight experiments that were designed to conceptually replicate past findings on CSTs and cognitive flexibility. Experiments 1–3 employ a previously used paradigm that involves imagined contact with a typical versus atypical Muslim and measure participants' cognitive reflection. In Experiments 4–6 a newly developed paradigm (imagined contact with a typical vs. atypical Chinese student) is used to examine whether it can be effective in boosting cognitive reflection. Finally, Experiments 7–8 employ yet another novel manipulation of imagined counter-stereotypical contact (with a male vs. female mechanic), which is subtler in nature than manipulations previously used. Across the eight experiments (total N = 1,137) the average sample-weighted meta-analytic effect was statistically non-significant and very small (d_+ = -0.13).

4.1 Exposure to Counter-stereotypes and Cognitive Flexibility

The experiments presented in this chapter aim to contribute to the triangulation of research findings on the effects of CSTs: The aim is to validate past findings through different research methods and multiple lines of evidence (Munafò & Smith, 2018). As reviewed in the preceding chapters, research on CSTs has typically employed paradigms that either (1) involved direct contact with counter-stereotypical people (e.g., Wilder, 1984), (2) required to form impressions of counter-stereotypical people (e.g., Hastie, Schroeder, & Weber, 1990), (3) used computer training techniques to "train stereotypes away" (e.g., Blair & Banaji, 1996), (4) involved mental imagery related to counter-stereotypical individuals (e.g., Blair, Ma, & Lenton, 2001), or (5) asked participants to generate counter-stereotypical exemplars (e.g., Vasiljevic & Crisp, 2013). Recently, a new paradigm has been presented that could be termed *imagined counter-stereotypical contact*. That is, Stathi, Crisp, and Hogg (2011), investigated whether imagined intergroup contact could generalise from being beneficial to one outgroup member to benefitting the outgroup as a whole. In the third of three experiments, the authors found that the positive effects of imagined contact most effectively generalised from one outgroup member to the outgroup as a whole when the imagined interaction involved an outgroup member who was typical rather than atypical. This finding is in line with Wilder's (1984) conjecture that a typical (versus atypical) outgroup member is likely to be more predictive

of their outgroup in general, and thus their behaviour is likely to be considered more informative. However, Stathi, Crisp, and Hogg (2011) did not design their experiments to test whether imagined counter-stereotypical contact can influence cognitive performance more generally, which is the aim of the present chapter.

Most recent research on CSTs and their cognitive downstream consequences (i.e., effects on cognitive flexibility and reflection) has used paradigms that either ask participants to form impressions of CST individuals or to generate CST exemplars themselves. In one line of research, Gocłowska, Crisp, and Labuschagne (2012) explored whether CSTs can generally boost flexible thinking that is unrelated to the social categories or the intergroup domain at hand. In the first of two experiments, the researchers found that thinking of a gender counter-stereotype (a female mechanic) versus a gender stereotype (a male mechanic) resulted in participants generating more cognitively flexible responses (the latter were measured as new names for pasta that participants generated). The effect was medium-to-large (Cohen's $d = 0.66$) and the researchers could rule out that this effect was due to changes in mood. In a second experiment, Gocłowska and colleagues conceptually replicated this finding with an alternative manipulation of CSTs and a different measure of cognitive flexibility. They asked participants to generate five counter-stereotypic (vs. stereotypic) category combinations and after that, to generate ideas for a themed night at the university nightclub as well as to prepare a poster advertising their idea. Again, Gocłowska and colleagues found that thinking of CSTs led to more general improvements in flexible and creative thought, an effect that was medium-sized (average Cohen's $d = 0.56$ for both creativity dependent variables

combined). To explain this finding, the authors proposed that thinking of CSTs activates a flexible thinking mindset or in other words, a *content-free processing orientation* that allows mental procedures exercised in one task to carry over into a new context (cf. Sassenberg & Moskovitz, 2005), although this potential mediator was not measured directly in this research.

In a related line of research, Vasiljevic and Crisp (2013) conceptually replicated the basic finding from Gocłowska, Crisp, and Labuschagne (2012) by employing similar manipulations of exposure to CSTs but alternative measures of cognitive flexibility. To measure a flexible thinking mindset, they used a self-report measure of need for cognitive closure (Experiment 1), the Stroop paradigm measuring participants' ability to inhibit stereotypic (i.e., dominant) associations (Experiment 2), and a measure of lateral thinking instructing participants to solve ten puzzles (Experiment 3). Results were in line with the authors' predictions, showing that participants who were asked to generate five counter-stereotypical exemplars outperformed participants who were asked to generate five stereotypical exemplars on the different cognitive flexibility tasks, effects that were medium-to-large in size (ranging from Cohen's $d = 0.51$ to $d = 0.84$). What is more, Vasiljevic and Crisp also examined whether thinking of CSTs could lead to a generalised reduction of prejudice towards multiple, unrelated outgroups (e.g., the elderly, disabled, asylum seekers, HIV patients). The results of three experiments——two lab and one field experiment——suggest that this was indeed the case. The researchers found that in addition to exposure to CSTs inducing a more flexible mindset, it also lowered prejudice towards multiple outgroups, increased commitment to democratic norms,

fostered egalitarian values, and increased trust towards outgroups in a setting characterised by a history of violent ethnic conflict (the Former Yugoslav Republic of Macedonia).

4.2 The Present Research

The present research aimed to conceptually replicate the key hypothesis postulated by the CPAG model and reported by Gocłowska et al. (2012) and Vasiljevic and Crisp (2013), which is that thinking of CSTs can create a flexible thinking mindset (operationalised as cognitive reflection). Specifically, the aim was to test this counter-stereotype-flexibility hypothesis using the imagined contact paradigm. Arguably, this paradigm differs from existing work on CSTs because it requires the mental simulation of social thought and action, thus potentially making the imagined scenario more personally meaningful, vivid, and complex (Crisp, Birtel & Meleady, 2011) as compared to mere exposure to counter-stereotypical information. This, in turn, should increase the power of the interventions (see also Lai et al., 2014, and Lai et al., 2016). To my knowledge, these experiments are the first to test whether imagined counter-stereotypical contact can activate a flexible thinking mindset. Based on the findings reported above, it was predicted that imagined contact with a counter-stereotypical versus stereotypical individual would boost cognitive reflection because it would induce a flexible thinking mindset. Experiments 1–3 tested this prediction by manipulating the stereotypicality of the target individual (typical vs. atypical Muslim) in the imagined contact scenario.

In Experiments 4–6, a second factor was added: The imagined contact took place at home versus abroad. This latter manipulation was added for exploratory purposes. For example, it is conceivable that imagined counter-stereotypical contact with an outgroup member abroad could be significantly more unexpected and impactful than imagined counter-stereotypical contact with an outgroup member at home. This is because people might expect to meet typical (rather than atypical) members of cultures when travelling abroad ("the outgroup homogeneity effect"; e.g., Quattrone & Jones, 1980). However, it is also possible that people might expect outgroup members to be more diverse than ingroup members ("the ingroup homogeneity effect"; e.g., Simon & Pettigrew, 1990). Because these are two competing predictions, no directional predictions were made with respect to imagined counter-stereotypical contact at home versus abroad. Rather, the goal was to explore whether the location (at home vs. abroad) would influence the effectiveness of the counter-stereotype interventions. Finally, Experiments 7–8 tested whether a subtle manipulation of imagined counter-stereotypical contact could have an effect on cognitive reflection. Note that across all experiments, the imagined contact scenarios had a positive tone in order to guard against the effects of a possible negative tone.

4.3 Experiments 1–3

4.3.1 Method

Participants. The minimum required sample size was calculated using G*Power 3.1 (Faul, Erdfelder, Lang & Buchner, 2007) based on the following assumptions. Given the findings reported by Gocłowska, Crisp, and Labuschagne (2012) and Vasiljevic and Crisp (2013), it was reasoned that the effect would be medium-to-large in size (Cohen's $d = 0.63$). Further, the alpha error probability was set at $\alpha = .05$ and a-priori power at 80%. Based on these assumptions, the minimum required total sample size was $N = 64$.

For Experiment 1, 79 University of Sheffield undergraduate students were recruited to take part in a laboratory experiment in return for course credit. To conceal the purpose of the study, a cover story in the consent form stated that the study would involve "a test of materials for a number of possible future studies". One participant guessed the purpose of the study upon completing it, therefore leaving a total sample size of $N = 78$ (78% female, 22% male, $M_{age} = 19.53$, $SD_{age} = 3.96$).

In Experiment 2, participants were recruited to take part in an online study via Prolific (www.prolific.ac), a platform that specialises in participant recruitment for research (Palan & Schitter, 2018; Peer, Brandimarte, Samat & Acquisti, 2017). In line with Prolific's minimum reward policy (which stipulates a minimum of £5/hour), participants were reimbursed with £1.70 for completing the study, which took approximately 20 minutes to complete. Contrary to Experiment 1, this study

was explicitly targeted towards English native speakers to ensure that instructions were understood. The cover story on the consent form was identical to Experiment 1. In addition, the following participant inclusion criteria were formulated a-priori: Participants who (1) self-identify as Muslim, (2) do not follow the instructions, (3) are not English native speakers, and (4) guess the purpose of the study would be excluded from analyses. From the 150 participants who completed the study, none of the participants guessed the purpose of the study, two participants reported to identify as Muslims, all participants followed the instructions, and all participants reported to be English native speakers. Hence, the final sample size was $N = 148$ (54% female, 46% male, $M_{age} = 24.03, SD_{age} = 6.56$).

In Experiment 3, participants were recruited via the University of Sheffield volunteers mailing list to take part in an online study. They were offered the chance to win one out of three shopping vouchers worth £10.00 each as a token of appreciation for their time. The cover story and participant inclusion criteria were identical to Experiment 2. From the 196 participants who completed the survey, none of the participants guessed the purpose of the study, nine participants reported to identify as Muslims, five participants did not follow the instructions, and 33 participants reported not to be English native speakers. Hence, they were excluded from analyses, leaving a final sample size of $N = 149$ (66% female, 34% male, $M_{age} = 27.66, SD_{age} = 11.70$).

Procedure. All experiments were programmed using Qualtrics survey software (www.qualtrics.com) and analysed using SPSS. The experimenter was always blind to condition. In Experiment 1, participants completed the experiment on computers

in individual laboratory cubicles. In Experiments 2 and 3, participants completed the experiment online and remotely in their own time. After providing informed consent, participants were randomly assigned to either the atypical or typical condition. All three experiments utilised an experimental manipulation taken from Experiment 3 in Stathi, Crisp, and Hogg (2011). That is, participants in the atypical condition read:

> "We would like you to take a minute to imagine yourself meeting a British Muslim stranger for the first time. Imagine that this person is not a typical British Muslim, he or she dresses in 'Western' clothes, drinks alcohol, eats pork and does not pray regularly. Imagine that the interaction is relaxed, positive and comfortable. Imagine that you learn about the life and experiences of your conversation partner."

In contrast, participants in the typical condition read:

> "We would like you to take a minute to imagine yourself meeting a British Muslim stranger for the first time. Imagine that this person is a typical British Muslim, he or she dresses in a traditional way, avoids alcohol, reads the Koran and prays five times a day. Imagine that the interaction is relaxed, positive and comfortable. Imagine that you learn about the life and experiences of your conversation partner."

Participants were given two minutes to imagine the interactions. After that, participants in both conditions were asked to describe what they had imagined in

as much detail as possible, which served to reinforce the manipulation. Participants were allowed to proceed to the next page as soon as they were satisfied with their response. This imagery description task was followed by the manipulation check, which consisted of the item, "Now, please think back to the imagery task: How typical was the Muslim you imagined of Muslims in general?" and was answered on a bipolar, 7-point scale ranging from 1 (*not at all*) to 7 (*very much*). After the manipulation check, participants completed ten puzzles (described in more detail below), which was followed by two items gauging participants' motivation and ability to engage with the imagination task. Specifically, the items were "During the imagery task I felt motivated to imagine the encounter with the Muslim" and "Throughout the imagery task I was able to imagine the interaction with the Muslim" and they were answered on bipolar, 7-point scales ranging from 1 (*not at all*) to 7 (*very much*). At the end of the experiment participants were asked whether they suspected that the study's purpose was something other than was stated, and if so, they were given the option to write down their thoughts regarding the purpose of the study. Lastly, participants answered some demographic questions (age, sex, first language), and then they were thanked and debriefed.

Dependent variable. In Experiments 1 and 2, the main dependent variable comprised a set of problems that have typically been used to assess flexible thinking, in particular the ability to switch from a heuristic to a more reflective mode of thinking. Ten problems were taken from West, Toplak, and Stanovich (2008) and Stanovich and West (2008). Although heuristic thinking tends to be adaptive in many cases (Gigerenzer, 2008), in some situations it can lead to poor decision making. As

such, the ability to override heuristic thinking by suppressing one's "impulsive" response in favour of more flexible thinking is crucial for successful problem-solving and decision-making (West et al., 2008). The ten problems used in the experiment can be found in Appendix A. As an example, consider this item:

> "Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which of the following options is likely? Circle one:
> (a) Linda is a bank teller.
> (b) Linda is a bank teller and is active in the feminist movement."

The first, intuitive answer that may come to mind is (b) because it appears to describe Linda more fittingly. However, (a) is, in fact, the correct answer. This is because (b) is a more specific case of (a), and thus (a) is more likely than (b). Realising that one's immediate response is incorrect is necessary to finding the correct solution to the puzzle. It was reasoned that participants who imagined contact with a counter-stereotypical individual would practise suppressing stereotypes and engage in individuating processing and thus, it was predicted that they would outperform participants in the stereotypical condition on the problem-solving task. The number of correct answers to the puzzles served as a measure of participants' performance.

In Experiment 3, a new dependent variable was used, but one that is conceptually related to the flexible thinking measure used in Experiments 1 and 2: The Cognitive Reflection Test (CRT; Frederick, 2005). The CRT is a widely used measure of

judgement and decision making and has been previously successfully used in research testing the effects of exposure to CSTs (e.g., Di Bella & Crisp, 2016). It comprises three items that measure the extent to which people engage in slow, reflective versus fast, heuristic thinking. For example, one CRT item reads:

> "In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?"

Whilst the intuitive answer to this question may be 24 days, the correct answer is in fact 47 days. As in Experiments 1 and 2, reaching the correct conclusion requires participants to (1) suppress the initial but incorrect response and (2) to reflect in order to determine the correct response. In other words, in order to perform well on this task participants are required to switch from the default heuristic processing mode to a more reflective, deliberative processing mode. To ensure that participants would complete the CRT promptly, they were given a maximum of two minutes to solve each of the three CRT questions.

4.3.2 Results

Experiment 1. An independent samples t-test confirmed the adequacy of the manipulation. Participants in the typical condition perceived the Muslim as significantly more typical of Muslims in general ($M = 5.55$, $SD = 1.22$) than participants in the atypical condition ($M = 2.95$, $SD = 1.41$), $t(76) = 8.68$, p(two-tailed)

$< .0005$, Cohen's $d = 1.97$, an effect that was very large (Cohen, 1988). Then a Univariate Between-Participants ANOVA was computed, with condition as the independent variable and the items measuring participants' motivation and ability to engage in the imagination task were as covariates in order to control for different levels of motivation and ability across conditions. The pattern of results was in the predicted direction, but the difference between conditions was not statistically significant. Participants in the atypical condition solved 4.63 problems on average ($SD = 2.08$) whereas participants in the typical condition solved 4.18 problems ($SD = 1.92$) on average, $F(1,75) = 2.01$, p(two-tailed) $= .16$, $d = 0.35$, which is a small effect (Cohen, 1988).

Experiment 2. An independent samples t-test confirmed that the manipulation was effective. Participants in the typical condition perceived the Muslim as significantly more typical of Muslims in general ($M = 5.00$, $SD = 1.49$) than participants in the atypical condition ($M = 3.12$, $SD = 1.34$), $t(146) = 8.09$, p(two-tailed) $< .0005$, Cohen's $d = 1.33$, an effect that was again large. A Univariate Between-Participants ANOVA was computed, with exactly the same parameters as in Experiment 1. The results revealed a statistically significant effect opposite to the predicted direction. Participants in the atypical condition performed significantly worse on the problem-solving task ($M = 4.85$, $SD = 2.02$) than participants in the typical condition ($M = 5.67$, $SD = 1.90$), $F(1,146) = 7.18$, p(two-tailed) $= .008$, $d = 0.48$, an effect that was medium-sized.

Experiment 3. Once again, an independent samples t-test confirmed the adequacy of the manipulation. Participants in the typical condition perceived the

Muslim as significantly more typical of Muslims in general ($M = 5.31$, $SD = 1.37$) than participants in the atypical condition ($M = 3.25$, $SD = 1.4$), $t(147) = 9.02$, p(two-tailed) $< .0005$, Cohen's $d = 1.50$, which is a large effect. A Univariate Between-Participants ANOVA was computed with exactly the same parameters as in Experiments 1 and 2. The pattern of results was in the predicted direction, but it was not statistically significant. Whereas participants in the atypical condition solved 1.46 CRT items on average ($SD = 1.19$), participants in the typical condition solved 1.28 items on average ($SD = 1.13$), with $F(1,148) = 0.66$, p(two-tailed) $= .42$, $d = 0.14$.

4.3.3 Discussion

The results of Experiments 1 through 3 do not provide support for the hypothesis that imagined contact with a counter-stereotypical (vs. stereotypical) individual promotes cognitive reflection. The results of Experiments 1 and 3 are not statistically significant and therefore inconclusive. The results of Experiment 2, in contrast, suggest that the experimental manipulation has a medium-sized effect on participants' cognitive flexibility *opposite* the predicted direction. Despite strong manipulations of counter-stereotypical experiences, which were identical across all three experiments, the results are inconsistent.

The reported patterns of results could have arisen for various reasons. First, Experiments 2 and 3 were conducted online, which means that there was significantly less control over the conditions in which participants completed the study.

Participants could have been multi-tasking while taking part in the study, or their environments may have distracted them. However, despite these possible circumstances inducing random error, it is clear from the manipulation checks that the experimental manipulations in each study were successful and powerful. Plus, any random error would have been equally likely in each condition, thus not posing any systematic problem. Indeed, recent research has confirmed that many lab-based experiments can successfully be replicated online (e.g., Crump, McDonnell, & Gureckis, 2013). In sum, it seems unlikely that the online nature of Experiments 2 and 3 had any major negative impact on the results.

Second, it is possible that the new measure of cognitive flexibility used in Experiment 3, the CRT, lacked sufficient sensitivity to capture the predicted effect. By virtue of only having three items (as compared to ten puzzles in Experiments 1 and 2), participants' problem-solving score only ranged from 0 (no correct answers) to 3 (all answers correct). This limited range of scores may not have been a sensitive enough measure to detect differences between conditions. This, however, does not explain why the pattern of results was completely reversed in Experiment 2, which had used a 10-item measure of flexible thinking. This raises an intriguing conundrum, which future research may be able to resolve.

A third interpretation of the inconsistent findings may be that the significance levels and effect sizes reported in the literature have been overestimated. Recently, the field-wide issue of potential publication bias has been extensively discussed (Ferguson & Heene, 2012; John, Loewenstein, & Prelec, 2012; Simmons, Nelson, & Simonsohn, 2011; but see also Fiedler & Schwarz, 2016). As a result of claims that

the field has been more likely to publish significant and novel findings, rather than replications or null effects, the true population effect size of a range of findings could have been overestimated. In this case, if the true population effect size of the effect of CSTs on flexibility is indeed smaller than expected, then the power analyses conducted for the present experiments may have been misguided and ultimately, by being under-powered, the studies may not have been able to detect effects that are small or medium-sized.

Finally, perhaps the most pessimistic interpretation is that the effect of CSTs on cognitive flexibility is a false-positive finding altogether. Albeit very unlikely, it is indeed possible that the published results do not represent a true effect but were due to chance——the very definition of p-values implies that 5% of statistically significant findings will be false positives, so it is important to keep this possibility in mind. Before any conclusions can be drawn, it is important to collect more evidence and synthesise it meta-analytically. Thus, in the following sections five more experiments are presented which will help to evaluate the reliability and generalisability of the effect of CSTs on cognitive flexibility.

4.4 Experiments 4–6

4.4.1 Method

Pilot study. New counter-stereotype manipulations were developed, which were pretested in a pilot study and compared with the manipulation used in Experi-

ments 1–3. The aim was determine which manipulation was most powerful, so it could be used for Experiments 4–6. Three new counter-stereotype manipulations were constructed: 1) traditional Muslim versus Muslim feminist, 2) typical versus atypical Chinese student, and 3) hetero- versus homosexual soldier. For the sake of comparison, the original manipulation (4) typical versus atypical Muslim was included in the pilot study (the verbatim phrasing of the new manipulations can be found in Appendix B). It was predicted that the atypical target individuals (i.e., Muslim feminist, atypical Chinese student, etc.) would be rated as significantly less typical of their respective social categories (i.e., Muslims, Chinese people, etc.) as compared to typical target individuals (i.e., traditional Muslims, typical Chinese student, etc.).

Because the manipulation checks in Experiments 1–3 suggested that the effect of condition on typicality ratings was very large (smallest Cohen's $d = 1.33$), the minimal required sample size for this study was $N = 64$ (eight participants per cell) according to a power analysis based on an alpha error probability of $\alpha = .05$ and a-priori power of 80%. A total of 89 participants were recruited via the Sheffield volunteers' mailing list (73% female, 27% male, $M_{age} = 21.93$, $SD_{age} = 6.65$) and randomly assigned to one of the eight conditions above. A Univariate Between-Participants ANOVA and planned contrasts were computed in order to compare the two conditions within each manipulation (see Figure 4.1).

The results of the pilot study revealed a highly statistically significant effect of condition on levels of typicality ratings, $F(7,88) = 11.80$, $p(\text{two-tailed}) < .0005$, $d = 1.44$, which is a very large effect. Planned contrasts showed that the 2^{nd} manip-

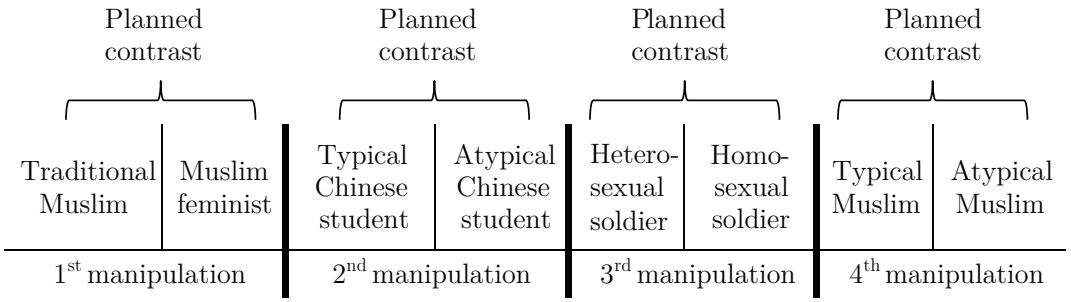

Figure 4.1: Analytic plan for the planned contrasts in the pilot study of Experiments 4–6.

ulation (typical vs. atypical Chinese student) was among the two most powerful manipulations. Participants in the typical Chinese student condition rated the target individual as significantly more typical of Chinese people in general ($M = 6.00$, $SD = .78$) than participants in the atypical Chinese student condition ($M = 3.36$, $SD = 1.50$), $t(21) = 3.41$, p(two-tailed) = .001, Cohen's $d = 1.45$, which is a large effect. According to the planned contrasts, the 4^{th} manipulation (typical vs. atypical Muslim, which was used in Experiments 1–3) was similarly powerful. Participants in the typical Muslim condition rated the target individual as significantly more typical ($M = 5.25$, $SD = 1.42$) than participants in the atypical Muslim condition ($M = 3.27$, $SD = 1.62$), $t(22) = 3.48$, p(two-tailed) = .001, Cohen's $d = 1.45$, which is an equally large effect. This was followed by the 1^{st} manipulation (traditional Muslim vs. Muslim feminist). Participants in the traditional Muslim condition rated the target individual as marginally more typical ($M = 5.10$, $SD = 1.29$) than participants in the Muslim feminist condition ($M = 4.00$, $SD = 1.34$), $t(20) = 1.80$,

p(two-tailed) = .076, Cohen's $d = 0.79$, which is a large effect. Finally, planned contrasts revealed that the 3^{rd} manipulation (typical vs. homosexual soldier) was the least powerful of the manipulations tested. Participants in the heterosexual soldier condition did not rate the target individual as significantly more typical ($M = 5.67$, $SD = 1.23$) than participants in the homosexual soldier condition ($M = 4.55$, $SD = 1.81$), $t(22) = 0.61$, p(two-tailed) = .54, Cohen's $d = 0.26$. In light of the results of the pilot study, the 2^{nd} manipulation (typical vs. atypical Chinese student) was chosen for Experiments 4–6. The results are depicted in Figure 4.2.

Participants. Given the uncertainty around the true effect size of the effect of CSTs on cognitive flexibility discussed above, it was not advisable to base the power analysis on the existing literature. Therefore, I did not aim to recruit a particular sample size in advance, but while keeping resource constraints in mind, I simply aimed to recruit as many participants as possible in all of the following experiments.

In Experiment 4, participants were recruited via the University of Sheffield volunteers' mailing list to take part in an online study. The participant reimbursement and the cover story were identical to Experiment 3. The participant inclusion criteria were the same as in Experiment 4, with one addition: If participants, for whatever reason, completed the study twice, then their duplicate submissions were removed. Of the 182 participants who completed the study, one of the participants guessed the purpose of the study, six participants did not follow the instructions, six participants reported not to be English native speakers, and one participant represented a duplicate. Hence, they were excluded from analyses, leaving a final

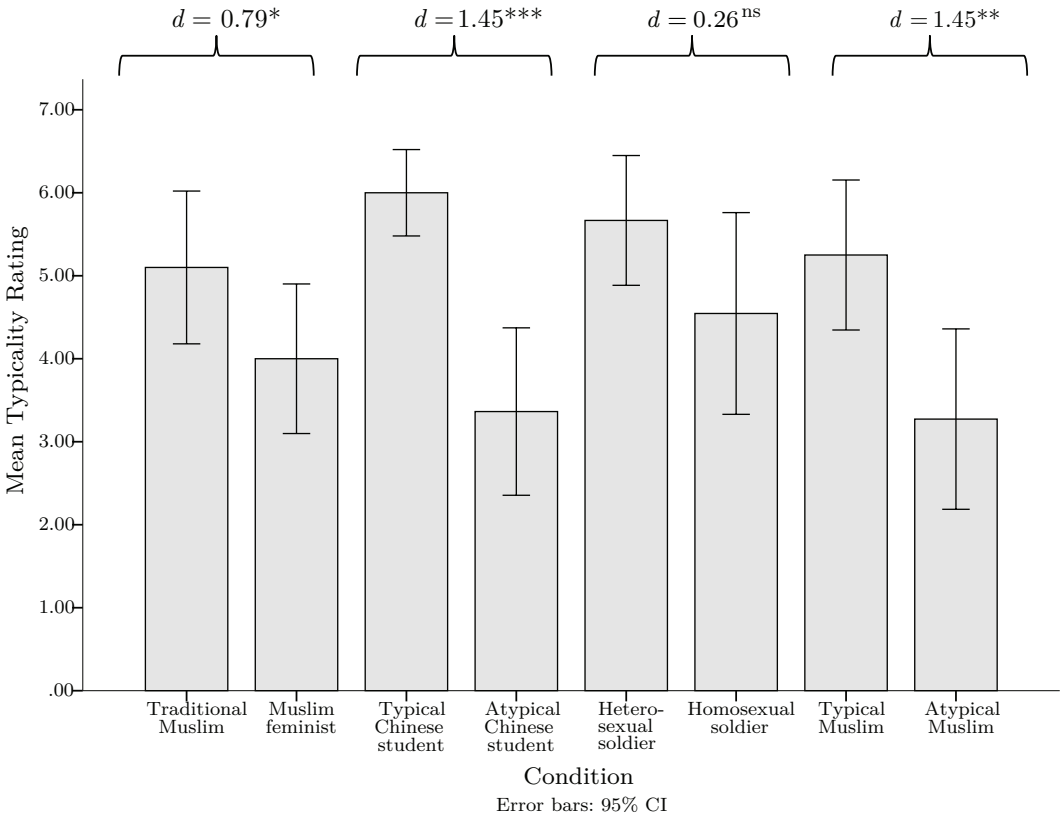

Figure 4.2: Results of Pilot Study 1. *** signifies $p < .0005$, ** signifies $p < .005$, * signifies $p < .05$, ^{ns} signifies statistical non-significance.

sample size of $N = 169$ (76% female, 24% male, $M_{age} = 23.76, SD_{age} = 9.49$).

In Experiment 5, participants were recruited via Amazon's Mechanical Turk (Buhrmester, Kwang & Gosling, 2011), an online platform comparable to Prolific (Peer et al., 2017). Because the imagined contact scenario was set in a university context, only students were recruited. Of 199 participants who completed the study, no participants guessed the purpose of the study, two participants did not follow

the instructions of the study, and two participants were not English native speakers and were thus removed. Additionally, six duplicate submissions were removed. Participants in this experiment were paid $1.50 each for their time and spent about 15 minutes completing it. Note that this participant sample primarily consisted of US American participants, which differs from the participant samples of previous experiments that involved mostly British citizens. This allowed to test for the effect of CSTs on cognitive flexibility in a different English-speaking country and culture, and to gauge the potential generalisability of the predicted effect. The final sample size was $N = 189$ (25% female, 75% male, $M_{age} = 25.43$, $SD_{age} = 5.63$).

In Experiment 6, participants were recruited in exactly the same way as the participants in Experiment 4. Of the 132 participants who completed the study, none of the participants guessed the purpose of the study, all participants followed the instructions, one participant reported not to be an English native speaker, and one participant represented a duplicate. Hence, they were excluded from analyses, leaving a final sample size of $N = 130$ (65% female, 35% male, $M_{age} = 28.08$, $SD_{age} = 11.91$).

Procedure. The procedure of Experiments 4–6 was identical to Experiment 2 except for the following change. In addition to manipulating stereotypicality, a further factor was added to explore whether the effects of counter-stereotypical imagery might have differential effects when imagined at home versus abroad. Thus, a 2 (typicality: stereotypical vs. counter-stereotypical) x 2 (location: at home vs. abroad) between participants factorial design was employed. The four conditions are presented in Table 4.1.

Table 4.1: A 2 (typicality: stereotypical vs. counter-stereotypical) x 2 (location: at home vs. abroad) between participants factorial design was used for Experiments 4–6. Information that differed across conditions is highlighted in bold.

	Typical	**Atypical**
At home	"**Imagine you are meeting a Chinese student at university for the first time.** After some time conversing with the student, you realise that this student is **harmony-oriented, compliant and diligent.** Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner."	"**Imagine you are meeting a Chinese student at university for the first time.** After some time conversing with the student, you realise that this student is **rebellious, individualistic and quite easy-going.** Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner."
Abroad	"**Imagine yourself traveling to China for the first time.** During your stay you get the chance to meet a Chinese student with a decent command of English. After some time conversing with the student, you realise that this student is **harmony-oriented, compliant and diligent.** Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner."	"**Imagine yourself traveling to China for the first time.** During your stay you get the chance to meet a Chinese student with a decent command of English. After some time conversing with the student, you realise that this student is **rebellious, individualistic and quite easy-going.** Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner."

Dependent variable. The dependent variable was identical to Experiment 2.

4.4.2 Results

Experiment 4. First, to check whether the manipulation was successful, a 2 x 2 Between-Participants ANOVA was conducted to test the main effects of the two independent variables (typicality, location) and their interaction on the typicality ratings, while controlling for participants' motivation and ability to engage in the imagination task. Typicality included two levels (typical vs. atypical imagined contact) and location included two levels (at home vs. abroad). As expected, the main effect of typicality on typicality ratings was statistically and practically significant, with $F(1,168) = 28.60$, p(two-tailed) $< .0005$, $d = 0.84$, indicating a large difference between the conditions. More precisely, participants in the typical condition perceived the Chinese student as more typical of Chinese students in general ($M = 4.80$, $SD = 1.55$) than participants in the atypical condition ($M = 3.52$, $SD = 1.55$). The main effect of location was not statistically or practically significant, with $F(1,168) = 1.88$, p(two-tailed) $= .173$, $d = 0.21$. More precisely, participants in the at home condition perceived the Chinese student as more typical of Chinese students in general ($M = 4.00$, $SD = 1.53$) than participants in the abroad condition ($M = 4.32$, $SD = 1.54$). Finally, the interaction effect was neither statistically nor practically significant, with $F(1,168) = 1.95$, p(two-tailed) $= .165$, $d = 0.22$.

Afterwards, to test the main hypothesis, a 2 x 2 Between-Participants ANOVA was conducted to test the main effects of the two independent variables (typicality, location) and their interaction on the number of problems solved, while controlling

for participants' motivation and ability to engage in the imagination task. As above, typicality and location included two levels, respectively. None of the effects were statistically or practically significant. The main effect of typicality yielded $F(1,168) = 1.00$, p(two-tailed) = .319, $d = 0.16$, indicating no statistically or practically significant difference between the typical ($M = 5.09$, $SD = 1.91$) and atypical condition ($M = 4.80$, $SD = 1.92$) in the number of problems solved. Further, the main effect of location yielded $F(1,168) = 1.89$, p(two-tailed) = .171, $d = 0.21$ indicating no statistically or practically significant difference between the at home ($M = 4.74$, $SD = 1.90$) and abroad condition ($M = 5.14$, $SD = 1.91$) in the number of problems solved. Finally, the interaction was also neither statistically or practically significant, with $F(1,168) = 0.17$, p(two-tailed) = .683, $d = 0.06$.

Experiment 5. First, to check whether the manipulation was successful, a 2 x 2 Between-Participants ANOVA was conducted to test the main effects of the two independent variables (typicality, location) and their interaction on the typicality ratings, while controlling for participants' motivation and ability to engage in the imagination task. Typicality included two levels (typical vs. atypical imagined contact) and location included two levels (at home vs. abroad). As expected, the main effect of typicality on typicality ratings was statistically and practically significant, with $F(1,188) = 33.49$, p(two-tailed) $< .0005$, $d = 0.86$, indicating a large difference between the conditions. More precisely, participants in the typical condition perceived the Chinese student as more typical of Chinese students in general ($M = 5.40$, $SD = 1.57$) than participants in the atypical condition ($M = 4.07$, $SD = 1.57$). The main effect of location was not statistically or practically

significant, with $F(1,188) = 1.27$, p(two-tailed) = .262, $d = 0.17$. More precisely, participants in the at home condition perceived the Chinese student as more typical of Chinese students in general ($M = 4.61$, $SD = 1.56$) than participants in the abroad condition ($M = 4.87$, $SD = 1.56$). Finally, the interaction effect was neither statistically nor practically significant, with $F(1,188) = 1.58$, p(two-tailed) = .21, $d = 0.19$.

Afterwards, to test the main hypothesis, a 2 x 2 Between-Participants ANOVA was conducted to test the main effects of the two independent variables (typicality, location) and their interaction on the number of problems solved, while controlling for participants' motivation and ability to engage in the imagination task. As above, typicality and location included two levels, respectively. None of the effects were statistically or practically significant. The main effect of typicality yielded $F(1,188) = 0.69$, p(two-tailed) = .408, $d = 0.13$, indicating no statistically or practically significant difference between the typical ($M = 5.15$, $SD = 2.26$) and atypical condition ($M = 4.87$, $SD = 2.27$) in the number of problems solved. Further, the main effect of location yielded $F(1,188) = 0.51$, p(two-tailed) = .478, $d = 0.11$ indicating no statistically or practically significant difference between the at home ($M = 5.13$, $SD = 2.25$) and abroad condition ($M = 4.89$, $SD = 2.26$) in the number of problems solved. Finally, the interaction was also neither statistically or practically significant, with $F(1,188) = 0.06$, p(two-tailed) = .815, $d < 0.01$.

Experiment 6. First, to check whether the manipulation was successful, a 2 x 2 Between-Participants ANOVA was conducted to test the main effects of the two independent variables (typicality, location) and their interaction on the typicality

ratings, while controlling for participants' motivation and ability to engage in the imagination task. Typicality included two levels (typical vs. atypical imagined contact) and location included two levels (at home vs. abroad). As expected, the main effect of typicality on typicality ratings was statistically and practically significant, with $F(1,129) = 25.89$, p(two-tailed) $< .0005$, $d = 0.92$, indicating a large difference between the conditions. More precisely, participants in the typical condition perceived the Chinese student as more typical of Chinese students in general ($M = 5.00$, $SD = 1.47$) than participants in the atypical condition ($M = 3.62$, $SD = 1.50$). The main effect of location was not statistically or practically significant, with $F(1,129) = 0.02$, p(two-tailed) $= .904$, $d < 0.01$. More precisely, participants in the at home condition perceived the Chinese student as more typical of Chinese students in general ($M = 4.31$, $SD = 1.52$) than participants in the abroad condition ($M = 4.27$, $SD = 1.47$). Finally, the interaction effect was neither statistically nor practically significant, with $F(1,129) = 0.003$, p(two-tailed) $= .953$, $d < 0.01$.

Afterwards, to test the main hypothesis, a 2 x 2 Between-Participants ANOVA was conducted to test the main effects of the two independent variables (typicality, location) and their interaction on the number of problems solved, while controlling for participants' motivation and ability to engage in the imagination task. As above, typicality and location included two levels, respectively. The main effect of typicality yielded $F(1,126) = 3.96$, p(two-tailed) $= .049$, $d = 0.36$, indicating a statistically or practically significant difference between the typical ($M = 5.68$, $SD = 1.86$) and atypical condition ($M = 5.01$, $SD = 1.89$) in the number of problems solved.

However, this effect was opposite the predicted direction. Further, the main effect of location yielded $F(1,126) = 0.38$, p(two-tailed) = .541, $d = 0.11$ indicating no statistically or practically significant difference between the at home ($M = 5.45$, $SD = 1.91$) and abroad condition ($M = 5.24$, $SD = 1.86$) in the number of problems solved. Finally, the interaction was neither statistically or practically significant, with $F(1,126) = 1.58$, p(two-tailed) = .212, $d = 0.23$.

4.4.3 Discussion

The results of Experiments 4–6 are similarly as inconsistent as the results of Experiments 1–3. The results of Experiments 4 and 5 were neither practically nor statistically significant. The results of Experiment 6, in contrast, suggest a difference in performance, again opposite the predicted direction: Participants in the typical condition outperformed participants in the atypical condition, an effect that was small. One limitation of Experiment 5 is that British participants were used to validate the materials in the pilot study, but US American participants were used in this experiment. Future research needs to ensure that study materials are validated within each relevant population. Lastly, the exploratory analysis exploring the role of the location of the imagined contact scenarios (at home vs. abroad) did not yield any statistically or practically significant interaction effects. Given the inconsistent results, it is difficult to draw any conclusions from the reported findings. Two further experiments will be presented that aimed to test whether a subtle manipulation of imagined counter-stereotypical contact could boost cognitive reflection. It was

reasoned that such a manipulation may be perceived as less threatening or aversive because it would minimise the extent to which the target individuals are perceived as deviants.

4.5 Experiments 7–8

4.5.1 Method

Pilot study. The new pilot study for Experiment 7 closely resembled the pilot study of Experiment 4. Overall, five possible CST interventions were pitted against each other. Three of the five interventions were identical to Pilot Study 1 (traditional Muslim vs. Muslim feminist, typical vs. homosexual soldier, and typical vs. atypical Muslim) and the remaining two interventions were newly constructed (male vs. female mechanic and typical vs. hippie lawyer). Again it was predicted that the atypical target individuals (i.e., Muslim feminist, female mechanic, etc.) would be rated as significantly less typical of their respective social categories (i.e., Muslims, mechanics, etc.) as compared to typical target individuals (i.e., traditional Muslims, male mechanics, etc.). The power analysis was identical to Pilot Study 1. A total of 322 participants were recruited via the Sheffield volunteers' mailing list and randomly assigned to one of the ten conditions above. Five participants were excluded because they indicated not to be English native speakers. The final sample size was thus $N = 317$ (69% female, 31% male, $M_{age} = 29.26, SD_{age} = 12.71$).

A Univariate Between-Participants ANOVA and planned contrasts were com-

puted in order to compare the two conditions within each manipulation. The results of this pilot study revealed a highly statistically and practically significant effect of condition on levels of typicality ratings, $F(9,315) = 13.41$, p(two-tailed) $< .0005$, $d = 1.25$. For the sake of brevity, only the results for the three most powerful interventions are reported. Planned contrasts showed that the 3rd manipulation (male vs. female mechanic) was the most powerful of the manipulations. Participants in the male mechanic condition rated the target individual as significantly more typical ($M = 5.35$, $SD = 1.53$) than participants in the female mechanic condition ($M = 2.70$, $SD = 1.65$), $t(69) = 7.17$, p(two-tailed) $< .0005$, Cohen's $d = 1.72$, which is a very large effect. According to the planned contrasts, the second most powerful manipulation was the 5th manipulation (typical vs. atypical Muslim). Participants in the typical Muslim condition rated the target individual as significantly more typical ($M = 5.31$, $SD = 1.38$) than participants in the atypical Muslim condition ($M = 3.00$, $SD = 1.10$), $t(72) = 3.48$, p(two-tailed) $< .0005$, Cohen's $d = 0.82$, which is a large effect as well. Finally, the third most powerful manipulation was the 3rd manipulation (typical vs. hippie lawyer). Participants in the typical lawyer condition rated the target individual as significantly more typical ($M = 4.61$, $SD = 1.67$) than participants in the hippie lawyer condition ($M = 3.63$, $SD = 2.09$), $t(65) = 3.04$, p(two-tailed) $= .004$, Cohen's $d = 0.75$, which is a large effect.

Participants. In Experiment 7, participants were recruited from the University of Sheffield psychology student participant pool and from the local Sheffield community to take part in a lab experiment. The participant reimbursement, the cover story, and the participant inclusion criteria were identical to the previous experiments.

Table 4.2: Instructions used for the new, subtle manipulation employed in Experiments 7–8. Information that differed across conditions is in bold; note that the only difference between conditions was the pronoun "he" versus "she".

	Typical	**Atypical**
At home	"Imagine yourself coming across a mechanic for the first time, who you talk to for a while. You find out that **he** is an auto mechanic working in a garage. Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner."	"Imagine yourself coming across a mechanic for the first time, who you talk to for a while. You find out that **she** is an auto mechanic working in a garage. Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner."

Of the 67 participants who completed the study, all participants met the inclusion criteria, hence leaving a final sample size of $N = 67$ (57% female, 43% male, $M_{age} = 31.58, SD_{age} = 15.75$).

In Experiment 8, participants were recruited via Prolific. Participants were reimbursed with £2.50 each for completing the study, which took approximately 25 minutes. Of the 220 participants who completed the study, three participants did not follow the instructions of the study and nine participants were not English native speakers. They were thus removed, leaving a final sample size of $N = 208$ (49% female, 51% male, $M_{age} = 22.21, SD_{age} = 4.35$).

Procedure. The procedure and dependent variable were identical to Experiments 1 and 2 except that the new manipulation (male vs. female mechanic) was employed (see Table 4.2).

4.5.2 Results

Experiment 7. An independent samples t-test confirmed the adequacy of the manipulation. Participants in the typical (male mechanic) condition perceived the mechanic as significantly more typical of mechanics in general ($M = 5.41$, $SD = 1.31$) than participants in the atypical (female mechanic) condition ($M = 2.97$, $SD = 1.59$), $t(65) = 6.88$, p(two-tailed) $< .0005$, Cohen's $d = 1.68$, an effect that was very large. Then a Univariate Between-Participants ANOVA was computed with the same parameters as Experiments 1–3. The results revealed no significant differences between conditions. Participants in the atypical condition did not perform better on the problem-solving task ($M = 5.15$, $SD = 1.66$) than participants in the typical condition ($M = 5.09$, $SD = 1.91$), $F(1,66) = 0.06$, p(two-tailed) $= .82$, $d = 0.06$.

Experiment 8. An independent samples t-test confirmed the adequacy of the manipulation. Participants in the typical (male mechanic) condition perceived the mechanic as significantly more typical of mechanics in general ($M = 5.47$, $SD = 1.34$) than participants in the atypical (female mechanic) condition ($M = 3.49$, $SD = 1.86$), $t(206) = 8.29$, p(two-tailed) $< .0005$, Cohen's $d = 1.18$, an effect that was large. Then a Univariate Between-Participants ANOVA was computed with the same parameters as Experiments 1–3. The results revealed that participants in the atypical condition performed slightly worse on the problem-solving task ($M = 5.32$, $SD = 2.08$) than participants in the typical condition ($M = 5.70$, $SD = 2.07$), $F(1,207) = 1.56$, p(two-tailed) $= .22$, $d = 0.19$, although this difference was not statistically significant.

4.5.3 Discussion

The results of Experiments 7 and 8 are largely in line with the results of Experiments 1 through 6——no statistically significant effects of the CST manipulation on cognitive reflection were detected that were in the predicted direction. In both Experiments 7 and 8, effect sizes were low, meaning that the results bear very little practical significance. In sum, the attempt to manipulate counter-stereotypicality in a rather subtle way did not seem to boost cognitive reflection.

4.6 Mini Meta-analysis

Because all eight experiments investigated the effect of exposure to CSTs on measures of cognitive reflection, a random-effects meta-analysis model was employed (using the Review Manager (RevMan) software; The Cochrane Collaboration, 2014) to estimate the average, sample-weighted effect of atypical (vs. typical) imagined contact on cognitive reflection.

The average effect of exposure to CSTs on cognitive reflection across the eight experiments (total N = 1,137) was d_+ = -0.13, $CI_{.95}$ [-.28, .02], which is a trend opposite the predicted direction. However, the fact that the 95% CI includes zero suggests that there is no statistically significant effect of the CST interventions on cognitive reflection (or at least not in these samples). This finding stands in contrast to previous research that has reported direct effects of CST interventions on cognitive performance (e.g., Gocłowska, Crisp, & Labuschagne, 2012; Prati, Vasiljevic,

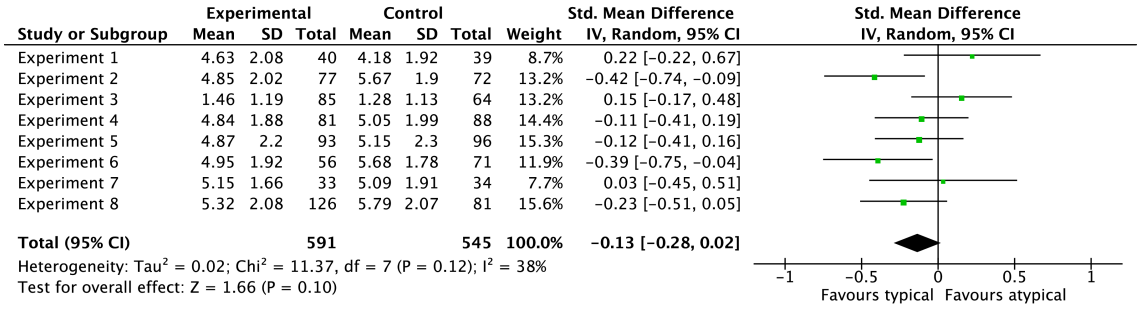

Figure 4.3: The results of the mini meta-analysis of Experiments 1–8.

et al., 2015b; Vasiljevic & Crisp, 2013). The results of the mini meta-analysis are depicted in Figure 4.3.

4.7 General Discussion

In eight experiments, the hypothesis that imagined counter-stereotypical contact could directly boost cognitive reflection was investigated. Despite strong manipulations of imagined counter-stereotypical contact, the experiments yielded rather inconclusive results. A mini meta-analysis revealed that on average, there was a trend opposite the predicted direction (d_+ = -0.13), which was a small effect (Cohen, 1988). Interestingly, it did not seem to matter whether imagined counter-stereotypical contact was imagined at home versus abroad, as all interaction effects were neither statistically nor practically significant.

4.7.1 Theoretical and Practical Implications

Given the inconclusive findings, only few direct implications can be drawn from the present experiments. First, in light of the reported results it appears that the large effect of CSTs on cognitive flexibility reported in past research (Gocłowska et al., 2012; Vasiljevic & Crisp, 2013), may be more constrained and smaller in size than previously thought. It may be that a very specific set of situational and/or individual conditions is required that allows for the effect to arise, or that the presence of certain circumstances easily undermines it (for example, fatigue, boredom, or motivated reasoning). If these conjectures are valid, then this has implications for models and theories specifying the effects of exposure to and contact with counter-stereotypical individuals. Specifically, in order to improve existing models, it will be necessary to detail the optimising and boundary conditions of the effect of CSTs on cognitive flexibility. In Chapter 5, three experiments will be presented that directly test one of the conditions postulated by the CPAG model, that is, the assumption that people have to be motivated to engage in inconsistency resolution in order for CSTs to affect broader cognitive functioning.

4.7.2 Limitations and Future Research

A myriad of reasons could explain the inconsistent pattern of findings. For example, it is possible that some of the experimental materials were lacking construct and/or content validity. For example, despite the fact that the problem-solving tasks closely resembled previously used tasks as part of counter-stereotype experiments

(e.g., Di Bella & Crisp, 2016), one cannot be confident that the problems adequately captured the relevant construct (cognitive reflection) because construct validity (e.g., convergent and divergent validity) has not been tested for the items used. Future research should explore the convergent and discriminant validity of the problem-solving measures employed here, and potentially also compute factor analyses in order to test whether the measures consist of one or multiple factors / dimensions.

In addition, it is possible that some of the conditions postulated by the CPAG model (see Section 2.4.2) were not met. Although in all experiments participants' motivation and ability to engage in imagined contact were controlled for, I did not measure whether participants did indeed engage in inconsistency resolution (i.e., stereotype suppression and generative thought). On the one hand, this initially did not seem necessary because previous research on CSTs only rarely measured these processes explicitly and yet reported large effects on cognitive performance. On the other hand, the fact that a measure of inconsistency resolution is lacking means that it is difficult to interpret why the predicted effect was not found. It may be that participants neither suppressed existing stereotypes, nor generated any new ideas to explain the stereotypical inconsistencies. Or it may be that they engaged in the former process (stereotype suppression), but not in the latter process (generative thought). Alternatively, it may be that the target individuals in the imagined scenarios were not sufficiently interesting or personally relevant to participants. In sum, it may be that the CST interventions that were employed in the experiments presented in this chapter did not sufficiently cognitively stimulate participants, and they therefore did not engage in inconsistency resolution. Future research should

measure whether or not participants engage in this process, which seems so critical to the effectiveness of CST manipulations.

Another possibility is that the CST manipulations employed in this research were not powerful enough to stimulate cognitive reflection. At first glance, this may contradict the fact that the manipulation checks always yielded large effect sizes. However, it could be that the manipulation checks themselves were not a good enough proxy for the effectiveness of the manipulation—the mere fact that participants in the atypical condition perceived the target individual as less typical than participants in the typical condition, does not necessarily imply that the manipulations were powerful. In future research, it may be helpful to develop more nuanced manipulation checks that can capture the "power" of the manipulation more adequately (e.g., ones that consist of multiple items rather than only one, or ones that use implicit rather than explicit measures).

Furthermore, it is important to consider that people can generally respond to CSTs in diverse ways. Existing research has distilled three broad categories for the way that people cognitively process CSTs, two of which lead to stereotype preservation (Plaks, Stroessner, Dweck & Sherman, 2001). First, upon encountering CSTs people may decrease their engagement with stereotype-inconsistent information and instead simply focus on stereotype-consistent information. Second, people may actively engage with stereotype-inconsistent information; however, at the same time try to "debunk or re-interpret it in a way that leaves the stereotype intact" (p. 876). It is presumed that these two strategies would contribute to the persistence of the stereotypes. Third, people may respond to CSTs with an open mind and by actively

engaging in the inconsistency resolution specified in the CPAG model. Regarding the paradigms used in this research, I expected the third strategy to be the one that is most likely employed by participants on average because I deemed the CSTs as non-threatening ("mild in intensity") and because the imagined contact scenarios were always positively valenced. However, unfortunately no measures were taken of any of these possible processes specified by Plaks et al. (2001), which means that no clear conclusions can be drawn in this regard. Future research is strongly advised to measure these processes directly.

Finally, as discussed in Section 4.3.3, publication bias in favour of significant and/or novel effects may have led to inflated estimates of effect sizes in the literature. As a result, it may be that the findings reported in the literature do not represent the true state of affairs with respect to the hypothesised counterstereotype-flexibility effect. In other words, the effect may either be much smaller than previously assumed, or not exist at all. Future research is required to evaluate the reliability, validity, and generalisability of the hypothesised effect of CSTs on cognitive reflection and flexibility.

4.7.3 Conclusion

This chapter explored whether the reported effect of CSTs on cognitive flexibility (e.g., Gocłowska et al., 2012; Vasiljevic & Crisp, 2013) could be conceptually replicated using alternative CST manipulations and a closely related dependent variable, cognitive reflection. The results of eight experiments——conducted both in the

laboratory and online—suggest that this effect is not as robust and generalisable as previously thought. Two out of eight experiments showed a statistically significant pattern opposite the predicted direction, and the remaining six experiments were inconclusive. A mini meta-analysis revealed a meta-analytic effect that is statistically non-significant and very small (d_+ = -0.13). In light of these findings, the next chapter delves more deeply into the question: What are the boundary conditions of the hypothesised effect of CSTs on cognitive flexibility, assuming it exists? Specifically, the premise that people have to be motivated to resolve stereotypical inconsistencies in order for CSTs to have effects on cognitive reflection will be examined in three experiments.

Chapter 5

The Effects of Counter-stereotype Interventions on Cognitive Reflection: The Moderating Role of Need for Cognition

---

Previous theorising and research has linked exposure to counter-stereotypical diversity (e.g., an Oxford-educated bricklayer) to enhanced cognitive performance and creativity. In this chapter, I examine the extent to which people's motivation to cognitively engage with the counter-stereotypical information (i.e., need for cognition, NFC) influences this effect. Across three experiments (N = 887) consistent support was found for the hypothesis that exposure to CSTs promotes cognitive reflection for people low in NFC (d_+ = 0.36), but not for people high in NFC (d_+ = -0.21).

The average direct effect of exposure to CSTs on cognitive reflection was statistically non-significant and small $d_+ = 0.08$. Further, moderated mediation analyses revealed that a psychological state characterised by surprise, perceived expectancy violation, and cognitive complexity mediated the moderated effect of condition on cognitive reflection.

5.1 Introduction

When the music legend Prince passed away, the New York Times honoured him by writing that he defied "narrow stereotypes about race and gender", thus "opening the minds of others" (The New York Times, 2016). Indeed, as the societal trends described in Chapters 1 and 2 suggest, contemporary societies increasingly bring people into contact with complex combinations of social, religious, and cultural identities (e.g., a female CEO, a Muslim hipster, a gay Catholic, a Harvard-educated carpenter). The question then, is how do people adapt to new forms of social and cultural diversity? And to what extent do individual differences play a role in how people respond to CSTs?

The present work is rooted in Crisp and Turner's (2011) CPAG model described in Chapter 2. One key assumption of the model is that people need to be motivated to resolve stereotypical inconsistencies in order for these to boost cognitive flexibility. For example, someone who meets an individual that challenges their stereotypes may wonder: "Why is this Muslim a hipster?", "How did this woman achieve the rank of a CEO?", or "What made this Harvard-educated man become a carpenter?" (Kunda,

1990). By seeking answers to these questions, judgments are less likely to be based on existing stereotypes and more likely on impressions of individuals. However, the hypothesis that perceivers need to be motivated to engage in *cognitive activity* in order for CSTs to promote cognitive performance has not been tested to date. More precisely, while some initial work has been done on the role of personal need for structure (PNS; i.e., preferences for the desired *outcome* of cognitive activity) in the link between exposure to CSTs and cognitive flexibility (e.g., Gocłowska, Baas, Crisp, & De Dreu, 2014), it is currently unknown whether need for cognition (i.e., preferences for the desired *amount* of cognitive activity) affects this link. In the present chapter I aimed to test this element of the CPAG model, that is, the role of individual differences in motivation in cognitive adaptation to counter-stereotypical diversity.

5.2 The Role of Need for Cognition in Exposure to Counter-stereotypes

In this research I set out to explore one theoretically important and practically meaningful potential moderator of the effects of exposure to CSTs on cognitive processing—namely, need for cognition (or NFC). NFC, also known as cognitive motivation or epistemic / intellectual curiosity (Mussel, 2010), is defined as "an individual's tendency to engage in and enjoy effortful cognitive activity" (Cacioppo, Petty, Feinstein & Jarvis, 1996). I suggest that NFC may influence the effect of

exposure to CSTs on cognitive reflection because CSTs are (by definition) surprising and novel (Prati et al., 2015a) and NFC is a marker of intellectual curiosity or the willingness to cognitively engage with novel stimuli (Litman & Spielberger, 2003; Mussel, 2010; von Stumm, Hell & Chamorro-Premuzic, 2011). Various studies have shown that NFC influences how people attend to, elaborate on, evaluate, and recall information (Peltier & Schibrowsky, 1994). Individual differences in NFC have been shown to predict complex problem solving and decision making (Nair & Ramnarayan, 2000), and intellectual curiosity has been identified as a key determinant of academic achievement (von Stumm et al., 2011). However, despite the evident importance of NFC in settings that are often characterised by social and cultural diversity, such as schools and companies, the hypothesis that NFC moderates the effects of exposure to diversity on cognitive performance remains untested.

5.3 Hypotheses

The present research tested three hypotheses. The first hypothesis was that exposure to CSTs will boost cognitive reflection (a main effect), as indicated by previous findings (Gocłowska et al., 2012; Prati et al., 2015b):

> Hypothesis 1 (main effect): Exposure to CSTs boosts cognitive reflection.

In addition, I expected that NFC would moderate the main effect of exposure to CSTs on cognitive reflection. More specifically, I hypothesised that exposure to

CSTs may spark interest in individuals low in NFC and motivate them to resolve the apparent stereotypical inconsistencies. In addition, I predicted that exposure to CSTs would make individuals low in NFC switch from a heuristic, category-based mode of processing to a systematic, individuating mode (Crisp & Meleady, 2012; Evans, 2008; Fiske & Neuberg, 1990; Strack & Deutsch, 2004). These predictions are consistent with Allen, Sherman, Conrey, and Stroessner (2009), who found that when people have low processing capacity and stereotypes are strong (e.g., a violent Black person, a warm and friendly woman), then they pay more attention to information that is inconsistent with their pre-existing stereotypes than information which is consistent. In contrast, for people high in NFC (who already engage in relatively systematic modes of processing by default) I predicted that exposure to CSTs might not be surprising enough to make them engage in even more systematic processing. Rather, I reasoned that exposure to CSTs might invoke a cognitively demanding inconsistency resolution process for people high in NFC, which could potentially be cognitively depleting (Hutter & Crisp, 2006). That is, people high in NFC run the risk of 'overthinking' or spending too much time trying to make

[1]Note that I initially entertained an alternative hypothesis, that is, that people high (rather than low) in NFC may be more likely to show cognitive reflection in response to exposure to CSTs. This is because their motivation to engage in cognitive activity may enhance the extent to which they are willing to expend resources in the face of expectancy-violating experiences (Gocłowska, Damian & Mor, 2017b; Leung & Chiu, 2008). In turn, this could mean that people high in NFC form more cross-cutting explanations for the inconsistent social categories, which may activate more distal cognitive associations and networks (Greenwald et al., 2002) and ultimately enhance cognitive reflection. However, I deemed that this alternative account was less likely to hold than

sense of CST individuals. In contrast, people low in NFC may find simpler ways to make sense of CSTs, leaving them with sufficient resources for subsequent cognitive reflection.[1]

> Hypothesis 2 (moderation effect): Exposure to CSTs boosts cognitive reflection for people low in NFC, but does not affect, or potentially even decreases, cognitive reflection for people high in NFC.

The third and final hypothesis concerns the psychological mechanism underlying the predicted effect of exposure to CSTs among people low in NFC. As discussed above, exposure to CSTs has previously been shown to elicit surprise (Prati et al., 2015a), and it is reasoned that people with a low NFC are more likely to experience surprise in response to CSTs. But research has also shown that surprise is not the only mediator that potentially explains the effects of exposure to CSTs. Specifically, Prati, Crisp, et al. (2015b) have demonstrated that exposure to CSTs not only prompts surprise, but also a perception that expectancies have been violated. Here it is suggested that in order for exposure to CSTs to improve performance, people need to view the relevant task as cognitively complex and stimulating. In other words, if people perceive the task as trivial and uninteresting, then they are unlikely to engage in it and therefore unlikely to reap the cognitive benefits following from

Hypothesis 2 (namely, that people low in NFC would be likely to benefit from exposure to CSTs), because people high in NFC may not be surprised by CSTs (or at least not as much as those low in NFC) and surprise has been shown to mediate the link between exposure to CSTs and (decreased) intergroup bias (Prati et al., 2015a).

exposure to CSTs. Taken together then, the final hypothesis extends Prati, Crisp, et al.'s (2015b) model to predict that a psychological state characterised by surprise, perceived expectancy violation, and cognitive complexity mediates the link between exposure to CSTs and enhanced cognitive performance for people low in NFC.

> Hypothesis 3 (moderated mediation effect): Exposure to CSTs boosts cognitive reflection for people low (but not high) in NFC because they experience a psychological state characterised by surprise, perceived expectancy violation, and cognitive complexity. This psychological state, in turn, increases cognitive reflection.

Three experiments tested the hypotheses by exposing participants to a variety of CSTs and subsequently measuring their cognitive reflection. I developed and validated two paradigms to solicit CST experiences and measured cognitive reflection using the 7-item Cognitive Reflection Test (CRT; Frederick, 2005; Toplak, West, & Stanovich, 2014). In order to test Hypothesis 2, I measured NFC using the 18-item NFC scale (Cacioppo, Petty & Kao, 1984). Hypothesis 1 and 2 were tested separately in each Experiment, but to maximise statistical power Hypothesis 3 (with respect to moderated mediation) was tested by pooling the data from all experiments.The experimental materials, data sets of the three experiments, and the *R* code used to run all analyses can be found on the Open Science Framework (https://goo.gl/CnYmsf).

5.4 Experiment 9

Pilot study. To manipulate exposure to CSTs, participants were asked to read a short paragraph, which described a CEO (Chief Executive Officer) named David. Participants in the control condition were asked to imagine that they read the following paragraph on the Internet: "David is a CEO. He's also a college graduate (Harvard), born and raised in the US, and happily married to his wife Linda". Participants in the experimental condition were asked to imagine that they read a slightly different paragraph about David: "David is a CEO. He's also a college dropout (Harvard), a Mexican immigrant, and happily married to his husband Michael." I established that the description of David was counter-stereotypical by recruiting 41 US American participants (16 female; $M_{age} = 31.51$, $SD_{age} = 11.53$) via Prolific and randomly assigning them to the two conditions described above. After reading the paragraph about David, participants were asked "To what extent do you feel surprised?" and "To what extent do you feel astonished?" on a scale from 0 (*not at all*) to 100 (*very much*). Next, to reinforce the manipulation, participants were instructed to imagine what David and his life were like and to describe (in as much detail as possible) their thoughts as to what characteristics he might possess. I checked that this manipulation was successful by asking participants to indicate their agreement with four statements: "David is a typical CEO" (reverse-coded), "Reading about David challenged some of my beliefs", "There isn't anything puzzling about David's life" (reverse-coded), and "Imagining David's life made me think 'outside the box'", again on a scale from 0 (Strongly disagree) to 100 (Strongly

agree). The manipulation check was followed by an attention check because it is often difficult to ascertain whether or not participants pay attention to the study materials (Oppenheimer, Meyvis & Davidenko, 2009), see Appendix C. I created a measure of counter-stereotypicality by calculating the mean of six items (i.e., the two items reflecting surprise and the four items reflecting counter-stereotypicality, α = .80). Lastly, participants were asked to indicate their sex, nationality, ethnicity, and English speaking ability, and were then thanked and debriefed.

It was predicted that David would be perceived as more counter-stereotypical in the experimental relative to the control condition. As expected, participants in the experimental condition did perceive David as significantly more counter-stereotypical (M = 48.09, SD = 16.67) than participants in the control condition (M = 16.24, SD = 9.95), with $t(29)$ = -7.20, p(two-tailed) < .001, Cohen's d = 2.34, which is a very large effect. In sum, these findings confirm the adequacy of the manipulation.

5.4.1 Method

Participants. Following previous findings exploring PNS as a moderator of the effect of CSTs on cognitive flexibility (Gocłowska & Crisp, 2013), I reasoned that the moderating effect of NFC would be medium-sized (d = .50). Power analysis, conducted using G*Power 3.1 (Faul et al., 2007), with an alpha of α = .05 suggested that N = 210 participants would provide 95% power to detect an effect of this magnitude. 397 participants were recruited via social media (www.reddit.com) and

Prolific to take part in an online experiment on "imagination and problem solving". Participants either volunteered their time or were compensated with GB£1.30 / US$1.80. The study duration was approximately 14 minutes. Participant inclusion criteria were determined a-priori: Only participants who a) took less than 30 minutes to complete the study, b) passed the attention check, and c) were not able to guess the purpose of the study were included in the analyses.[2] The final sample size was $N = 315$ participants (177 male, 134 female, 3 other, 1 prefer not to say; $M_{age} = 29.87$, $SD_{age} = 10.57$; 86% US American nationality, 14% other).

Procedure. The experiment comprised three parts and participants completed all tasks online using the survey software Qualtrics (www.qualtrics.com). Part 1 was identical to the pre-test in that participants were randomly assigned to imagine a stereotypical or a counter-stereotypical CEO named David and then asked to indicate their surprise and astonishment like in the pilot study. Next, to reinforce the manipulation, participants were instructed to imagine what David and his life could be like and to describe what characteristics he might possess.

[2]These inclusion criteria were applied to all experiments in this chapter. Removing participants who took longer than 30 minutes to complete the task was important because it helped establish confidence that the participants completed the studies with as few distractions as possible. This was particularly important for Experiment 10 because it was not possible to set a time limit in this experiment (unlike in Experiments 9 and 11, which were conducted using an online platform that allowed this facility). Because participation in this experiment was completely voluntary (i.e., without a guaranteed incentive), many participants took an unusually long time to complete Experiment 10 ($M = 37.50$ minutes, $SD = 102.21$ minutes), which may have diminished possible effects of the intervention.

In part 2, participants' cognitive reflection was measured using the 7-item version of the CRT (Toplak, West & Stanovich, 2014). Recall that the items are designed such that an incorrect solution to each of the seven questions initially comes to mind. Cognitive reflection is demonstrated when the incorrect response is overridden and, upon further reflection, the correct solution is determined. For example, one item states that "Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?" The intuitive, but incorrect, answer is "30"; while the correct answer is "29". Participants were presented with seven such problems in a counter-balanced order and were given up to two minutes to solve each problem. They were automatically redirected to the next page when the time was up. If they solved a problem in less than two minutes, then they were allowed to proceed. The problem-solving task was followed by the remaining four manipulation check items and the attention check, which were identical to the pre-test. The manipulation check items had good internal consistency ($\alpha = .82$).

In part 3 of the experiment, participants were asked to indicate whether they were suspicious at any point that the researchers were investigating something other than what was stated, and if so, they were asked to describe what they thought the real purpose of the study was. Next, participants indicated their age, sex, sexual orientation, nationality, ethnicity, and English speaking ability. Finally, participants completed the 18-item NFC scale. Upon completing these questionnaires, participants were thanked for participating and debriefed.

Analytic approach. The data were analysed using moderated regression analyses with the pequod package in the programming language *R* (Mirisola & Seta, 2016).

Conditions were contrast coded as –1 (control) and +1 (experimental) and a mean score reflecting NFC was computed by averaging the 18 items (reverse-coded where appropriate, $\alpha = .95$). Because 86% of the sample was US American but the remaining 14% had other nationality, I also report the moderated regression analyses for the US sample only. This served to check whether cross-national differences in participants might have influenced the effect.

5.4.2 Results and Discussion

Manipulation check. To check the adequacy of the CST manipulation, the mean CST score (i.e., the index of the manipulation check items) was regressed on condition, NFC, and their interaction. The independent variables were centered prior to computing the interaction term. As expected, there was a main effect of condition (b = 12.11, $t(310)$ = 12.12, p(two-tailed) < .001), such that participants in the experimental condition perceived David as more counter-stereotypical than participants in the control condition. The effect of NFC (b = -0.65, $t(310)$ = -0.94, p(two-tailed) = .35) and the interaction term (b = 0.41, $t(310)$ = 0.60, p(two-tailed) = .55) were not statistically significant. Thus, the manipulation of counter-stereotypicality was successful, regardless of participants' level of NFC.

Randomisation check. Prior to testing Hypothesis 2——that NFC moderates the effect of exposure to CSTs on cognitive reflection——it was checked whether NFC differed across conditions. This is because NFC had been measured as part of the same experimental session and, although unlikely, participants' responses to the

measure of NFC may have been affected by the experimental manipulation. Welch's t-test revealed that NFC did not significantly differ across conditions, $t(312) = 0.08$, p(two-tailed) = .93, Cohen's $d = 0.009$, suggesting that the manipulation did not affect NFC scores.

The effects of condition, NFC, and their interaction on cognitive reflection. To test Hypothesis 2, the above analysis was repeated, but this time regressing the number of correctly solved CRT-items on condition, NFC, and their interaction. The analyses revealed no main effect of condition ($b = 0.01$, $t(311) = 0.11$, p(two-tailed) = .92), but a statistically significant main effect of NFC ($b = 0.39$, $t(311) = 4.81$, p(two-tailed) $< .001$), such that participants high in NFC consistently outperformed participants low in NFC on the cognitive reflection task. This is not surprising, as previous research has demonstrated that NFC predicts cognitive performance (Cacioppo et al., 1996).

The main effect of NFC was, however, qualified by a marginally statistically significant two-way interaction between condition and NFC (b = -0.15, $t(311)$ = -1.87, p(two-tailed) = .063). To understand the nature of the interaction, the effect of condition (experimental vs. control) on cognitive reflection at different levels of NFC was inspected using simple slopes analysis (Aiken & West, 1991). "Low NFC" was defined as 1 SD below the mean and "high NFC" as 1 SD above the mean. The analysis revealed a pattern of results in the predicted direction, however while it was not statistically significant, it has some practical significance. A positive slope was found for participants low in NFC ($b = 0.24$, $t(311) = 1.40$, p(two-tailed) = .16) and a negative slope for participants high in NFC (b = -0.21, $t(311)$ = -1.24, p(two-

tailed) = .21 (see Figure 5.1). Although neither of these slopes reached statistical significance, it is important to note that the performance gap between participants low versus high in NFC was reduced in the experimental (vs. control) condition.[3]

Discussion. The findings of Experiment 9 provide no support for Hypothesis 1 (that exposure to CSTs generally boosts cognitive reflection, regardless of individual differences in motivation). However, the findings do provide preliminary and very tentative evidence in support of Hypothesis 2. That is, the marginally significant interaction effect indicates that NFC may moderate the effects of exposure to CSTs such that it may benefit people low but not high in NFC, although these trends did not reach statistical significance. Interestingly, the effects became somewhat more pronounced when the analysis was re-run with the US American sample only, which potentially suggests that cross-national differences may have weakened the effect. Given that the critical interaction in Experiment 9 was only marginally significant, Experiment 10 aimed to replicate these findings with an alternative manipulation of exposure to CSTs. A measure of mood was also included (namely,

[3]Running the analysis with the US sample yielded a more clear-cut pattern of results. Again, there was no effect of the experimental condition on CRT-performance (b = 0.05, $t(267)$ = 0.36, p(two-tailed) = .72) and there was a highly statistically significant main effect of NFC on CRT-performance (b = 0.39, $t(267)$ = 4.75, p(two-tailed) < .001). The two-way interaction between condition and NFC became slightly more pronounced (b = -0.16, $t(267)$ = -1.93, p(two-tailed) = .055). Simple slopes analyses revealed a positive trend for the experimental condition on the performance of participants low in NFC (b = 0.28, $t(267)$ = 1.62, p(two-tailed) = .11), but no effect of the experimental condition on the performance of participants high in NFC (b = -0.20, $t(267)$ = -1.11, p(two-tailed) = .27).

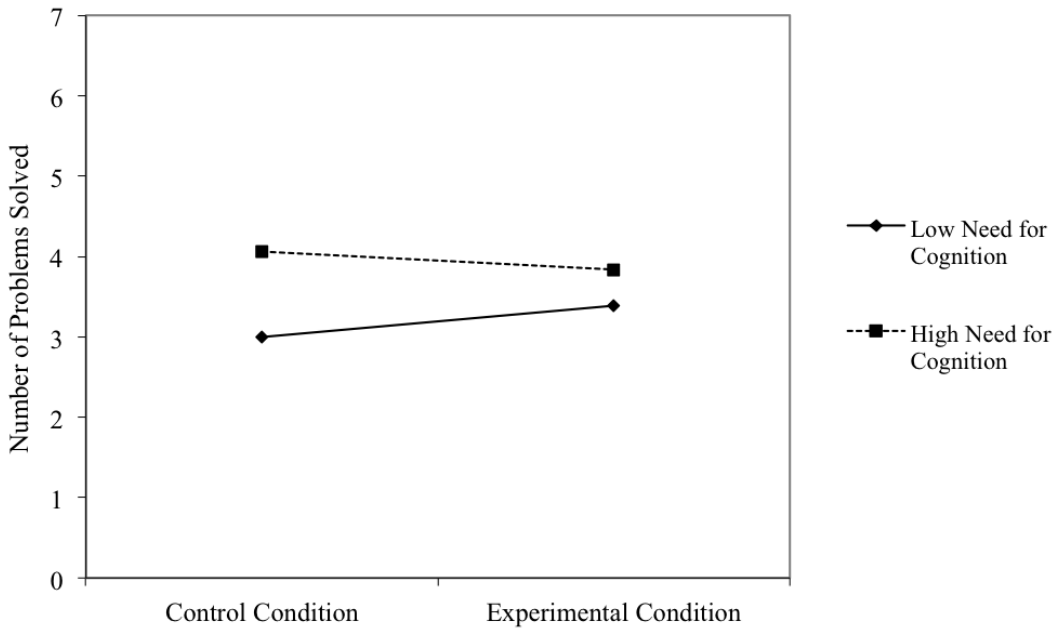

Figure 5.1: Cognitive reflection as a function of exposure to CSTs at different levels of NFC (Experiment 9).

the brief mood introspection scale; Mayer & Gaschke, 1988) because both positive and negative moods have previously been linked to enhanced cognitive and creative performance (Baas, De Dreu & Nijstad, 2008; Cheng, Leung & Wu, 2011; Isen, Daubman & Nowicki, 1987). Including the brief mood introspection scale allowed to investigate whether the effect of exposure to CSTs on cognitive performance holds when controlling for different mood states, and thus to rule out mood as an explanation for the hypothesised effect.

5.5 Experiment 10

Pilot study. As before, participants were asked to read a paragraph, but this time describing a person named Mary. Participants in the control condition were asked to imagine that they read the following paragraph on the Internet: "Mary is a secondary school teacher (married, two children), a university graduate (English literature), and UK native. Mary has a positive outlook on life." Participants in the experimental condition were asked to imagine reading the following paragraph instead: "Mary is a political leader (remarried, two children), a scientist (quantum physics), and a Polish immigrant. Mary has a positive outlook on life." A pre-test was used to establish the counter-stereotypicality of these new stimulus materials. Specifically, 51 British participants (25 female; $M_{age} = 34.06$, $SD_{age} = 10.15$) were recruited via the crowdsourcing platform Prolific and randomly assigned to imagine Mary the (stereotypical) female teacher or Mary the (counter-stereotypical) female political leader. After reading the paragraph about Mary, participants were asked how surprised and astonished they felt and were instructed to imagine what Mary and her life could be like. Following this task, participants indicated their agreement to four statements: "Mary is a typical woman" (reverse-coded), "Reading about Mary challenged some of my beliefs", "There isn't anything puzzling about Mary's life" (reverse-coded), and "Imagining Mary's life made me think 'outside the box'", all on a scale from 0 (Strongly disagree) to 100 (Strongly agree). A composite measure of counter-stereotypicality was again created by calculating the mean of the six items, that is, the items measuring surprise and astonishment and the four

items measuring counter-stereotypicality. The internal consistency of these items was acceptable (α = .69). Lastly, participants were asked to indicate their sex, age, nationality, ethnicity, and English speaking ability, and were then thanked and debriefed.

As expected, participants in the experimental condition perceived Mary as significantly more counter-stereotypical (M = 37.29, SD = 17.35) than participants in the control condition (M = 26.47, SD = 10.14), $t(40)$ = -2.73, p(two-tailed) = .009, Cohen's d = 0.76, which was a large effect.

5.5.1 Method

Participants. Based on the calculation of statistical power in Experiment 9, I again aimed to recruit a minimum of 210 participants to Experiment 10. 616 participants were recruited via a university mailing list at a UK university to take part in an online experiment on "imagination and problem solving". All participants who completed the experiment were entered into a prize draw to win one of two GB£50.00 shopping vouchers or one of five GB£20.00 shopping vouchers. The attention check and participant inclusion criteria were identical to Experiment 9. The final sample consisted of 302 participants (90 male, 206 female, 3 other, 3 prefer not to say; M_{age} = 24.21, SD_{age} = 8.12; 81% British nationality, 19% other nationality).

Procedure. The procedure and materials were identical to Experiment 9 except for the new manipulation (i.e., Mary the female teacher vs. political leader, rather

than David the CEO) and the addition of the brief mood introspection scale (Mayer & Gaschke, 1988) after the attention check. To examine whether exposure to CSTs altered mood states related to cognitive reflection, I computed variables representing positive activating moods (lively, happy, peppy, loving, caring, and active; $\alpha = .74$), negative activating moods (jittery, nervous, fed up, gloomy, grouchy, and sad; $\alpha = .80$), positive deactivating moods (content and calm; $\alpha = .52$), and negative deactivating moods (tired and drowsy; $\alpha = .66$), which were measured using a 7-point Likert scale ranging from 1 (*not at all*) to 7 (*very much*). The manipulation check items had acceptable internal consistency ($\alpha = .69$).

5.5.2 Results and Discussion

Manipulation check. To check the adequacy of the CST manipulation, hierarchical regression was used to examine the effect of condition on the mean CST score while controlling for the different mood states. At step 1, condition, NFC, and their interaction term were entered as independent variables and the mean CST score as the dependent variable. This produced a highly significant main effect of condition on the mean CST score ($b = 6.10$, $t(283) = 7.53$, p(two-tailed) $< .001$), such that participants in the experimental condition viewed Mary as more counter-stereotypical than participants in the control condition. There was no significant effect of NFC on the mean CST score ($b = 0.38$, $t(283) = 0.47$, p(two-tailed) $= .64$), but there was a marginally significant interaction effect between NFC and condition on the mean CST score ($b = -1.44$, $t(283) = -1.82$, p(two-tailed) $= .07$). In step

2, the four mood variables were added as control variables to the above regression equation. This produced a pattern of results similar to step 1. The main effect of condition on the mean CST score remained highly significant (b = 6.14, $t(273)$ = 7.43, p(two-tailed) < .001), such that participants in the experimental condition viewed Mary as more counter-stereotypical than participants in the control condition. Like in step 1, there was no significant effect of NFC on the mean CST score (b = 0.56, $t(273)$ = 0.69, p(two-tailed) = .49), but the interaction effect between NFC and condition on the mean CST score approached significance (b = -1.23, $t(273)$ = -1.51, p(two-tailed) = .13). In addition, there was a marginally significant effect of negative activating moods on the mean CST score (b = 3.68, $t(273)$ = 1.86, p(two-tailed) = .06). These results are interpreted as suggesting that a) the CST manipulation was successful because of the highly significant main effect of condition on the mean CST score and b) that participants' mood did play a role in the effects of exposure to CSTs on perceptions of counter-stereotypicality. To control for the effects of mood on the dependent variables, the four mood variables were entered as covariates to the regression models in all subsequent analyses.

Randomisation check. A Welch Two Sample t-test revealed that NFC did not significantly differ between the conditions ($t(277)$ = -0.87, p(two-tailed) = .38, Cohen's d = .10) suggesting that the randomisation to the experimental versus control condition was successful.

The effects of condition, NFC, and their interaction on cognitive reflection. Next, to test Hypothesis 2, I repeated the above moderated regression analysis, but this time regressing the number of correctly solved CRT-items on condition, NFC, and

their interaction, with and without controlling for the four different mood states. In step 1, condition, NFC, and their interaction term were entered as independent variables and the CRT performance at the dependent variable. The analyses revealed a trend for the experimental condition to influence CRT-performance (b = 0.20, $t(284)$ = 1.68, p(two-tailed) = .09) such that participants in the experimental condition outperformed participants in the control condition. There was also a highly significant main effect of NFC (b = 0.47, $t(284)$ = 3.99, p(two-tailed) < .001) such that participants high in NFC outperformed participants low in NFC. The main effects were, however, qualified by a statistically significant two-way interaction between condition and NFC (b = -0.25, $t(284)$ = -2.10, p(two-tailed) = .04). Simple slopes revealed a positive effect of the experimental condition on performance for participants low in NFC (b = 0.46, $t(284)$ = 2.67, p(two-tailed) = .008), but no effect of the experimental condition on performance for participants high in NFC (b = -0.05, $t(284)$ = -0.30, p(two-tailed) = .76). In step 2, the mood variables were added as covariates to the regression model. Again, the analyses revealed a trend for experimental condition to influence CRT-performance (b = 0.20, $t(274)$ = 1.68, p(two-tailed) = .10) such that participants in the experimental condition outperformed participants in the control condition. Again, there was also a highly significant main effect of NFC (b = 0.48, $t(274)$ = 3.98, p(two-tailed) < .001) such that participants high in NFC outperformed participants low in NFC. The main effects were, however, qualified by a significant two-way interaction between condition and NFC (b = -0.28, $t(274)$ = -2.29, p(two-tailed) = .02). Simple slopes analyses revealed a positive effect of the experimental condition on the performance

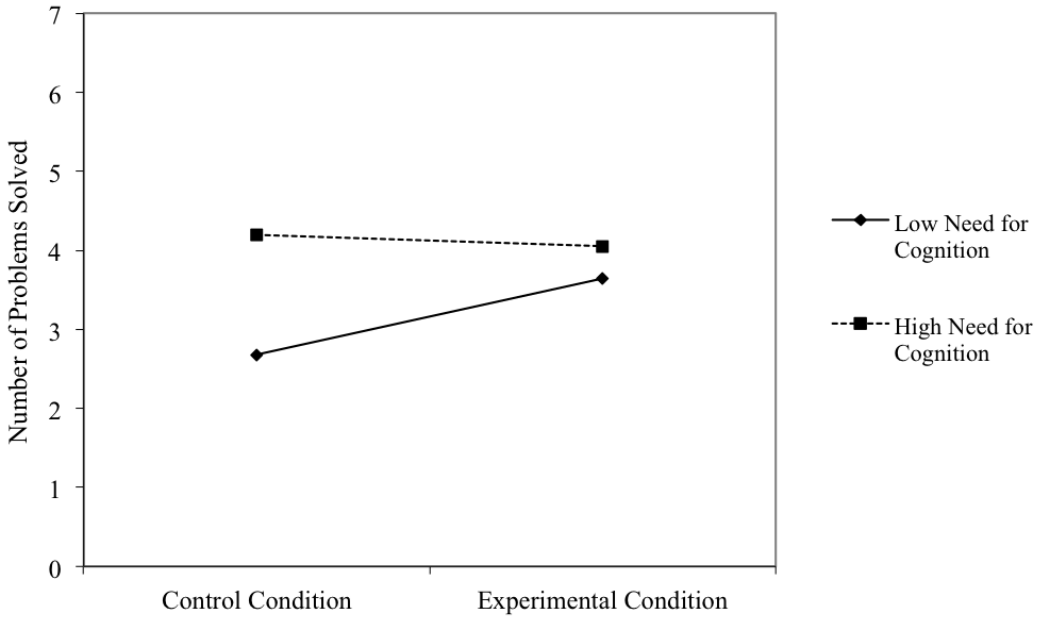

Figure 5.2: Cognitive reflection as a function of exposure to CSTs at different levels of NFC while controlling for mood (Experiment 10).

of participants low in NFC (b = 0.49, $t(274)$ = 2.79, p(two-tailed) = .006), but no effect of the experimental condition on the performance of participants high in NFC (b = -0.08, $t(274)$ = -0.45, p(two-tailed) = .65). These results suggest that including the mood variables as covariates improved the explanatory power of the regression model. Moreover, it is noteworthy that the performance gap between participants low versus high in NFC was again reduced in the experimental (vs. control) condition. Figure 5.2 illustrates these findings.

Discussion. The findings of Experiment 10 provide some evidence in support of Hypothesis 1 (which states that exposure to CSTs generally boosts cognitive

reflection, regardless of individual differences), although the main effect of condition was only marginally significant. The findings do, however, provide further evidence for Hypothesis 2, which is that exposure to CSTs benefits people low, but not high in NFC. However, one limitation with Experiments 9 and 10 is that counter-stereotypicality was manipulated and cognitive reflection and NFC were measured in the same experimental session. Although the randomisation checks showed that NFC did not differ across conditions, the next aim was to provide a more rigorous test of the hypotheses in Experiment 11, which separated the measure of NFC from the experimental manipulation by a week.

5.6 Experiment 11

5.6.1 Method

Participants. Once again I aimed to recruit a minimum of 210 participants. 344 participants were recruited via Prolific to take part in an online experiment on "imagination and problem solving" in return for GB£1.60. The study duration was approximately 14 minutes. The attention check and participant inclusion criteria were identical to Experiments 9 and 10. The final sample consisted of 270 participants (96 male, 171 female, 2 other, 1 prefer not to say; $M_{age} = 31.59$, $SD_{age} = 10.77$; 99% British, 1% other).

Procedure and materials. The procedure and materials were identical to Experiment 10 except for the following changes. In part 1 of the experiment, participants

answered the NFC scale and demographic questions (sex, age, nationality, ethnicity, English language ability), whereas part 2 was administered one week later and involved the CST manipulation and CRT. Experiment 11 also incorporated two attention checks. The first attention check was the same as in Experiments 9 and 10 and was placed in part 1 of the experiment. The second attention check was placed in part 2 of the experiment (see Appendix D).

5.6.2 Results and Discussion

Manipulation check. To check the adequacy of the CST manipulation, I again used hierarchical regression to examine the effect of condition on the mean CST score while controlling for the different mood states. At step 1, condition, NFC, and their interaction term were entered as independent variables and the mean CST score at the dependent variable. This produced a highly significant main effect of condition on the mean CST score (b = 8.35, $t(256)$ = 8.88, p(two-tailed) < .001), such that participants in the experimental condition viewed Mary as more counter-stereotypical than participants in the control condition. There was no significant effect of NFC on the mean CST score (b = -0.83, $t(256)$ = -1.00, p(two-tailed) = .32), but there was a marginally significant interaction effect between NFC and condition on the mean CST score (b = -1.50, $t(256)$ = -1.82, p(two-tailed) = .07). In step 2, the four mood variables were added as control variables to the above regression equation. This produced a pattern of results similar to step 1. The main effect of condition on the mean CST score remained highly significant (b =

8.40, $t(248) = 9.04$, p(two-tailed) $< .001$), such that participants in the experimental condition viewed Mary as more counter-stereotypical than participants in the control condition. Like in step 1, there was no significant effect of NFC on the mean CST score (b = -0.87, $t(248)$ = -1.08, p(two-tailed) = .28), and the interaction effect between NFC and condition on the mean CST score was not significant either (b = -1.12, $t(248)$ = -1.35, p(two-tailed) = .18). However, there was a statistically significant effect of positive activating moods on the mean CST score (b = 6.75, $t(248) = 3.08$, p(two-tailed) = .002). The highly significant main effect of condition on the CST score suggests that the CST manipulation was successful. Again, the effects of mood on the dependent variables were controlled for in all subsequent analyses.

The effect of condition, NFC, and their interaction on cognitive reflection. To test the hypotheses, I repeated the above moderated regression analysis, but this time regressing the number of correctly solved CRT-items on condition, NFC, and their interaction, while controlling for the four different mood states. At step 1, condition, NFC, and their interaction term were entered as independent variables and CRT performance as the dependent variable. This produced a non-significant main effect of condition on CRT performance (b = 0.03, $t(258)$ = 0.19, p(two-tailed) = .85) and a highly significant main effect of NFC on CRT performance (b = 0.53, $t(258)$ = 4.18, p(two-tailed) $< .001$). The main effects were qualified by a significant interaction effect between NFC and condition on CRT performance (b = -0.27, $t(258)$ = -2.44, p(two-tailed) = .016). In step 2, the four mood variables were added as control variables to the above regression equation. This produced a

pattern of results similar to step 1. The main effect of condition on CRT performance remained non-significant (b = 0.01, $t(250)$ = 0.08, p(two-tailed) = .94) and the effect of NFC on CRT performance (b = 0.53, $t(250)$ = 4.77, p(two-tailed) < .001) remained highly significant. Like in step 1, the interaction effect between NFC and condition on CRT performance was statistically significant (b = -0.30, $t(250)$ = -2.67, p(two-tailed) = .008). Simple slopes analyses revealed a positive effect of condition on CRT performance for participants low in NFC (b = 0.36, $t(250)$ = 1.95, p(two-tailed) = .05) and a negative (albeit marginally significant) effect of condition on CRT performance for participants high in NFC (b = -0.34, $t(250)$ = -1.88, p(two-tailed) = .06). Again, the performance gap between participants low versus high in NFC was reduced in the experimental (vs. control) condition. The results are depicted in Figure 5.3.

Discussion. The findings of Experiment 11 yield no evidence in support of Hypothesis 1 (namely, that exposure to CSTs generally boosts cognitive reflection), but they provide further evidence in support of Hypothesis 2 (namely, that people low but not high in NFC benefit from exposure to CSTs). In the next two sections, I combine the insights from the three experiments meta-analytically and conduct a moderated mediation analysis to test Hypothesis 3, which concerns the potential mechanism that underlies the effect of exposure to CSTs on cognitive reflection.

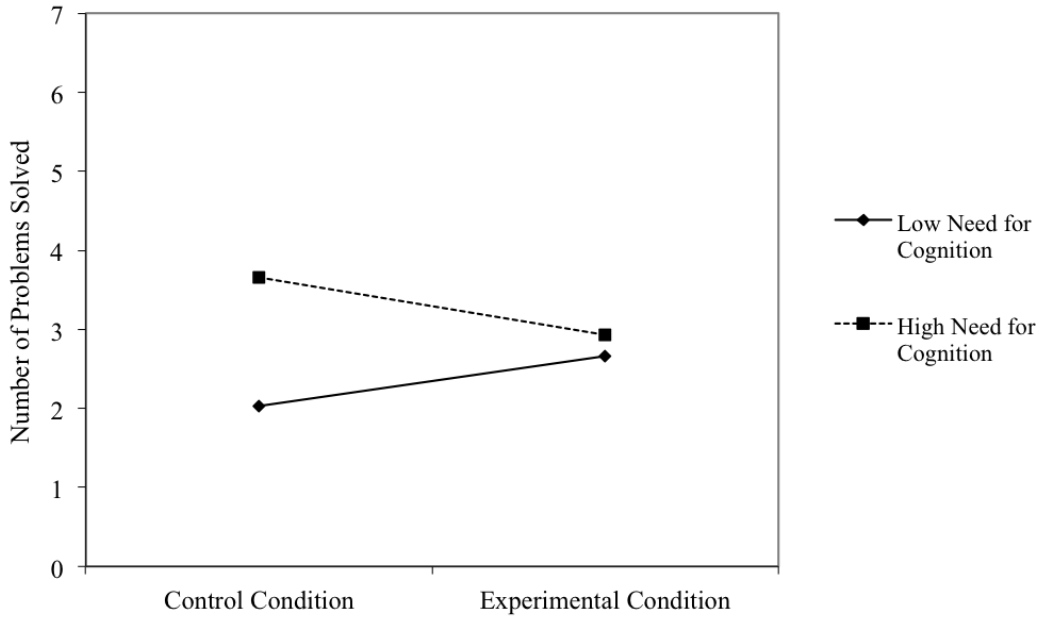

Figure 5.3: Cognitive reflection as a function of exposure to CSTs at different levels of NFC, controlling for mood (Experiment 11).

5.7 Meta Summary of Effect Sizes across the Experiments

Because all three experiments investigated the effect of exposure to CSTs on measures of cognitive reflection, I employed a fixed-effects meta-analysis model (using the metafor package in *R*; Viechtbauer, 2010) to estimate the average effect of exposure to CSTs on cognitive reflection. Specifically, I computed the sample-weighted (main) effects of condition on cognitive reflection, respectively, across the sample as a whole and also the effect of condition among participants low versus

high in NFC separately.

The average effect of exposure to CSTs on cognitive reflection across the three experiments was $d_+ = 0.08$, $CI_{.95}$ [-0.06, 0.22]. The fact that the 95% CI includes zero suggests that exposure to CSTs does not generally boost cognitive reflection (or at least not in the present samples), thus yielding no evidence in support of Hypothesis 1. This finding stands in contrast to previous research that has reported main effects of exposure to CSTs on cognitive performance (Gocłowska et al., 2012; Prati et al., 2015a; Vasiljevic & Crisp, 2013).

Recall, however, that it was also hypothesised that the effect of exposure to CSTs would differ (or be moderated) by levels of NFC. In support of this idea, on average, exposure to CSTs had a small positive effect on the cognitive reflection of participants low in NFC, $d_+ = 0.36$, $CI_{.95}$ [0.16, 0.56], while exposure to CSTs had a small negative effect on the cognitive reflection of participants high in NFC, $d_+ =$ -0.21, $CI_{.95}$ [-0.40 -0.01]. This finding supports Hypothesis 2, that exposure to CSTs boosts cognitive reflection for people low in NFC, but reduces cognitive reflection of people high in NFC.

5.8 Moderated Mediation Analysis across the Experiments

In order to maximise statistical power to detect a moderated mediation effect, I pooled the data from all three experiments to test whether a psychological state

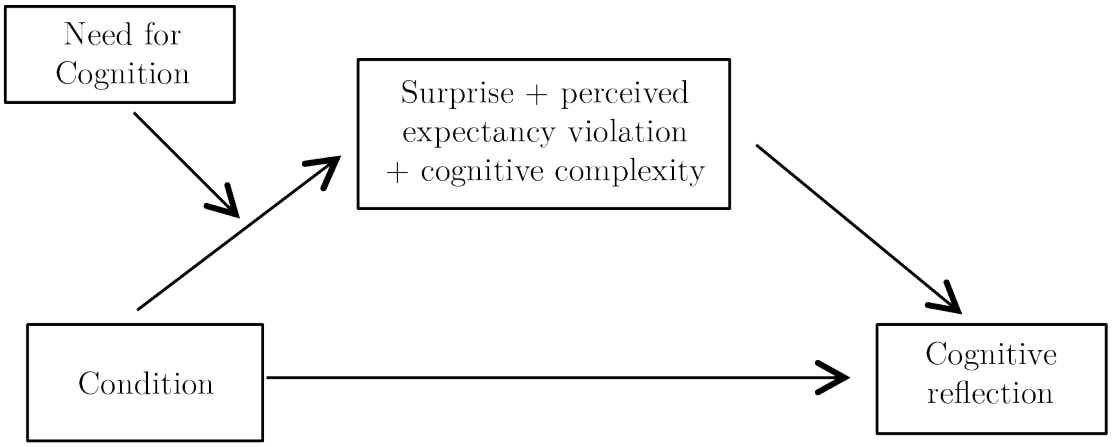

Figure 5.4: Hypothesised moderated mediation model.

characterised by surprise, perceived expectancy violation, and cognitive complexity mediated the moderated effect of condition on cognitive reflection. This moderated-mediation prediction was tested using the lavaan package in *R* (Rosseel, 2012). The model is depicted in Figure 5.4.

First, I constructed a variable representing the proposed mediator by averaging measures of the three component processes (surprise, perceived expectancy violation, and cognitive complexity; $\alpha = .59$). The exact items constituting the new composite mediator variable can be found in Appendix E. I then ran a moderated mediation analysis testing the model in Figure 5.4. The data were treated as the population and 1,000 bootstrap samples were drawn (with replacement) to create 95% bias-corrected confidence intervals. Like in Experiments 10 and 11, the analysis controlled for positive and negative activating and deactivating moods. The final sample size was

N = 536.

The results revealed a marginally significant interaction effect between condition and NFC on the composite mediator (b = -1.24, SE = .65, p(two-tailed) = .06), as well as a significant effect of the composite mediator on CRT performance (b = 0.02, SE = .006, p(two-tailed) = .01). Decomposing the interaction effect revealed that the indirect effects at the different levels of NFC (–1SD and +1SD, respectively) were statistically significant but varied in size. Specifically, exposure to CSTs prompted participants low in NFC to experience a psychological state characterised by surprise, perceived expectancy violation, and cognitive complexity, which led to improved CRT performance (b = 0.12, SE = .05, p(two-tailed) = .02). Interestingly, the same effect was found for people high in NFC, however it was smaller in size (b = 0.08, SE = .04, p(two-tailed) = .03). This supports Hypothesis 3, which stated that people low in NFC would be more likely to experience a psychological state characterised by surprise, perceived expectancy violation, and cognitively complexity than people high in NFC, which in turn would be more likely to boost their cognitive performance.

5.9 General Discussion

Three experiments examined how exposure to CSTs affects cognitive reflection among participants who are low versus high in NFC. As predicted, participants low in NFC performed better on the CRT following exposure to CSTs than did participants low in NFC who were not exposed to CSTs. Across the three experi-

ments, the average effect of exposure to CSTs among participants low in NFC was small-to-medium in magnitude (d_+ = .36). Interestingly, exposure to CSTs also influenced the performance of participants high in NFC on the Cognitive Reflection Test. However, unlike participants low in NFC, the cognitive performance of participants high in NFC tended to decrease following exposure to CSTs——an effect that was, on average, small in magnitude (d_+ = -.23). Taken together, these findings provide converging evidence that the effects of interventions based on exposure to CSTs depend on, or are moderated by, individual differences in cognitive motivation. Furthermore, in line with my expectations, a psychological state characterised by surprise, perceived expectancy violation, and cognitive complexity mediated the effect of exposure to CSTs on cognitive reflection for people low in NFC. Contrary to my expectations, the same psychological state also mediated the effect of exposure to CSTs on cognitive reflection for people high in NFC, albeit to a lesser degree.

5.9.1 Limitations and Future Directions

The findings with respect to participants low in NFC were as anticipated on the basis of theory and past research. However, it is less obvious why participants high in NFC show decreased performance after exposure to CSTs. It is also important to consider how this effect might be reconciled with the finding that they also experienced more surprise, expectancy violation, and cognitive complexity, like the participants low in NFC, who showed improved cognitive performance. One possibility is that exposure to CSTs also arouses different, potentially antagonistic

component processes among people high in NFC. Specifically, exposure to CSTs may trigger cognitive depletion and / or self-regulatory fatigue for people high in NFC, in addition to the psychological challenge response that is captured the experience of surprise and expectancy violation. Thus, people high in NFC may experience competing psychological processes in response to CSTs——on the one hand, they feel somewhat intrigued and interested, but on the other hand they also feel cognitively depleted. Further research is needed to test this idea.

In addition, limitations of the present research concern internal and external validity. Regarding internal validity, it is important to note that stereotype content slightly differed across conditions, and thus potentially could have confounded the results. For example, in Experiments 10 and 11 Mary was a schoolteacher in the control condition, but a political leader in the experimental condition. Whilst both professional occupations require a certain level of expertise and competencies, it is conceivable that one of the two——e.g., Mary the political leader——may have been perceived as higher in status or power, which may (or may not) have caused changes in cognitive performance independent of (counter-)stereotypicality (e.g., by making participants feel more threatened by the CST target individual).

5.9.2 Theoretical and Practical Implications

There has been a surprising dearth of research on exposure to diversity and how its effects on cognitive outcomes differ between individuals. By identifying one moderating variable——namely, NFC——and how it may influence the effect of exposure

to CSTs on cognitive reflection, the present research represents an important step towards closing this gap. This chapter demonstrates that a simple "one size fits all" explanation of how exposure to CSTs influences performance is overly simplistic. Failing to consider individual differences in NFC in the effects of exposure to CSTs may, for example, unintentionally give rise to adverse consequences for people high in NFC. The practical implications are that both researchers and practitioners need to consider NFC as a moderating variable when designing CST interventions.

The research reported in this chapter is consistent with research on personal need for structure (PNS) as a moderator of the effect of exposure to CSTs on creativity (Gocłowska et al., 2014; Gocłowska & Crisp, 2013). According to Gocłowska and Crisp (2013), people high in PNS seek to organise information in relatively simple ways and therefore dislike experiences that challenge their mental representations. In contrast, people low in PNS approach tasks in a more open-minded manner and are less inclined to over-generalise, which predisposes them to embrace inconsistencies. In line with this reasoning, Gocłowska and Crisp (2013) found that exposure to a CST (a female mechanic) only enhanced creative performance among individuals low in PNS. Similarly, Gocłowska et al. (2014) found that exposure to information that violates social schemas (e.g., a female mechanic, an astronaut on the beach, an Eskimo in a desert, a hockey player on a football field) promoted creativity among people low in PNS, but impeded creativity among people high in PNS. In line with these earlier findings, the present work suggests that participants low in NFC (akin to those low in PNS) have more "headroom"—that is, more room to open up and be cognitively stimulated—than those high in NFC. This, in turn,

may help people low in NFC engage with the counter-stereotypical information and may ultimately allow for improved cognitive performance. Future research needs to test this prediction directly.

On a theoretical level, the present research extends previous work on exposure to CSTs and its effects on affective, motivational, and cognitive outcomes. Specifically, the three experiments reported in this chapter suggest that exposure to CSTs can sometimes be remarkably powerful, which has theoretical implications for models of cognitive adaptation to diversity (Crisp & Turner, 2011; Gocłowska et al., 2017b). That is, it appears that high levels of cognitive motivation may not be required in order for people to engage with CSTs, but instead a low level of cognitive motivation appears sufficient to enable cognitive reflection after exposure to CSTs. However, an important caveat is that too much cognitive motivation can potentially backfire.

5.10 Conclusion

The role of NFC / intellectual curiosity in the effect of exposure to counter-stereotypical diversity on cognitive performance has been relatively neglected in social psychological research to date, despite its importance for professional and educational settings. Three experiments (N = 887) support the hypothesis that exposure to CSTs has a positive effect on the cognitive reflection of people low in NFC, but a negative effect on the cognitive reflection of people high in NFC. Taken together, this research contributes to a more nuanced understanding of the effects of exposure to CSTs on cognitive reflection, which in turn could help to maximise the

gains and minimise the pains of diversity (Galinsky et al., 2015). The next chapter turns to a longitudinal analysis of exposure to CSTs and examines the effects it may have on cognitive flexibility, intergroup bias, concern for discrimination, and epistemic unfreezing.

Chapter 6

Repeated Exposure to Counter-stereotypes

In this chapter, I review and discuss whether exposure to CSTs can lead to a sustained change in cognitive flexibility and reductions in intergroup bias. I present and validate a new paradigm that can be used to experimentally study the longitudinal effects of exposure to CSTs and that allows for control of differences in stereotype content. I also present the results of a preregistered longitudinal experiment that examined the effects of exposure to CSTs on cognitive flexibility, explicit bias, concern for discrimination, and "epistemic unfreezing" (a psychological process characterised by curiosity and open-mindedness) over the course of 14 days. Finally, I discuss the findings of the longitudinal experiment and reflect on their implications for research and theories on psychological adaptation to social and cultural diversity.

6.1 Introduction

In January 2015, a campaign called 'This Girl Can' was launched by Sport England, with the aim to encourage women to play sports. The campaign was based on a survey conducted in 2014 suggesting that women were significantly less likely to exercise than men: Specifically, it was found that two million fewer women than men regularly played sports, although over 75% of women reported a desire to exercise more (Sport England, 2014). With slogans such as "Sweating like a pig, feeling like a fox" or "A kick right in the stereotypes", the campaign aimed to prompt a change in attitudes and gender stereotypes and help boost women's confidence when playing sports. About £10 million has been invested into the campaign, and follow-up research suggests that 250,000 women have started exercising as a result of it (Sport England, 2016). This campaign represents a great example of an intervention in which people were repeatedly exposed to CSTs (via TV campaigns, Youtube videos, and content shared on social media[1]). The effectiveness of the campaign was assessed using the Active People Survey, which monitors the amount of sports people play (Sport England, 2017). However, it appears that the effects of the campaign were not evaluated rigorously because no randomised controlled trials were conducted to directly test the intervention's effects. Therefore, it is unclear how many of the women began exercising as a result of the intervention, or as a result of other factors (e.g., a general change in attitudes towards sports).

The aim of the research described in this chapter was to develop and implement a

[1]Examples of This Girl Can imagery can be found here: https://goo.gl/k6EqqA

new intervention that draws on the power of CSTs and that could be administered repeatedly over time. For a long time, researchers have lamented the scarcity of studies testing the longitudinal effects of diversity interventions (e.g., Pettigrew, 2008). Albeit scarce, some examples of such interventions do exist. For instance, Wilder and Thompson (1980) investigated whether repeated, positive contact with outgroup members would be more effective in decreasing intergroup bias as compared to a single contact session among US American college students. The authors found that two contact sessions were more beneficial than only one——indeed, as the amount of contact with outgroup members increased, intergroup bias decreased. However, this only held true for measures that pertained to the outgroup situation at hand (e.g., attitudes towards specific group members and the willingness to grant rewards to them), and the effect did not generalise towards the outgroup as a whole (e.g., when attitudes towards the outgroup in general were measured). Wilder and Thompson speculated that changes in general attitudes towards the outgroup might have occurred relatively slowly (giving rise to a "sleeper effect") and could therefore not have been detected. Alternatively, the outgroup members in the experiment may have been considered atypical and therefore not representative of the outgroup as a whole (i.e., subtyped). A brief, open-ended follow-up with a random selection of the participants lent support to the latter explanation——participants regarded the experimental situation as highly unusual, which may have prevented more positive attitudes from generalising to the outgroup. A third, methodological limitation might also have contributed to the finding that the effect of the intervention did not generalise towards the outgroup as a whole: It is possible that the study was

underpowered and thus could not detect effects that were smaller. After all, there were only 23 participants per condition in this experiment—based on power analysis (assuming $\alpha = 0.05$ and power of 80%) this only allows sufficient power to detect effects that are large. Nevertheless, this experiment provides some evidence that the frequency of contact with outgroup members can matter to some extent: If more frequent contact helps to reduce intergroup bias (at least towards the outgroup members in the given situation), then one first step towards improved intergroup relations is made, which can be further built on.

6.2 The Longitudinal Effects of Exposure to Counter-stereotypes

Fast forward to contemporary psychological science, and one observes a discussion around the longevity of intervention effects. In 2012, Devine, Forscher, Austin, and Cox published research testing a multi-faceted intervention that aimed to reduce implicit race bias over time. The intervention was based on three premises: In order to combat implicit bias, it is important for people to (1) develop awareness of their biases, (2) show concern about the effects of these biases, and (3) actively use strategies that can help to diminish the biases. In a 12-week longitudinal experiment, the researchers found that participants in the intervention group showed a decrease in implicit racial bias as compared to participants in the control group. What is more, participants who showed concern about racial discrimination and

participants who reported having used the strategies (e.g., stereotype replacement, counter-stereotypic imaging, perspective taking), showed the greatest reductions in implicit bias. Moreover, intervention (but not control) participants showed an increase for concern in racial discrimination over time and heightened personal awareness of bias. Interestingly, however, the intervention did not affect levels of explicit bias, which was measured using the Attitudes Towards Blacks scale (including items such as "Generally, Blacks are not as smart as Whites" or "It would not bother me if my new roommate was Black"; Brigham, 1993).

Forscher, Mitamura, Dix, Cox, and Devine (2017) set out to replicate these findings in a well-powered longitudinal experiment. In line with previous results, the researchers found a change in concern for discrimination that persisted for two weeks after the intervention. However, in contrast to the findings of Devine et al. (2012), the intervention was not found to change implicit bias for participants in the intervention versus control condition——instead, participants in all conditions showed a decrease in implicit bias over time. Forscher and colleagues also tried to disentangle the effects of the different prejudice reduction strategies (e.g., counter-stereotypical imagery) on the dependent variables. The researchers found evidence suggesting that counter-stereotypical imagery decreased concern for discrimination (b = -0.34). One interpretation of this finding is that emphasising examples of outgroup members who seem unaffected by discrimination may have resulted in an ironic negative effect on concern. Taken together, Forscher et al. concluded that because "the present study had a much larger sample size than [Devine et al., 2012] (...) [and] as long as the present study is a fair test of the effects of the habit-

breaking intervention, it is more likely that the effects (...) reported by Devine and colleagues are false positives than that the present results are false negatives" (p. 143).

Forscher et al.'s research conveys at least two important insights for diversity interventions and diversity science more generally. First, it underlines the importance of replications: They are a cornerstone of science, ensuring that resources are invested into programmes of research and interventions that yield reliable outcomes. Second, it suggests that even a highly intensive, involving longitudinal intervention to combat prejudice may fall short of changing both explicit and implicit bias. In addition, Forscher et al.'s findings raise a number of questions and conundrums, such as: What can most effectively be done to reduce intergroup bias? What are the most critical psychological mechanisms, contextual variables, and the most promising strategies for improving intergroup relations? The next section reviews work by Lai et al. (2016), which helps shed some light on potential answers.

Lai et al. (2016) evaluated the effectiveness of a range of diversity interventions across time. The authors aimed to answer: Can interventions that have been shown to immediately reduce implicit racial preferences (e.g., imagining a vivid counter-stereotypical scenario, forming implementation intentions) induce lasting change in implicit bias? Nine interventions were tested (eight real ones and one sham intervention) in two studies (total N = 6,321) in US American universities. While this research primarily focused on implicit bias as a dependent variable (as measured by the IAT), the researchers also included measures of explicit racial bias (e.g., rating warmth for White and Black people), and a measure of political preferences

(i.e., support for affirmative action helping Black people). The results showed that all of the interventions reduced implicit bias immediately, generally replicating the findings of Lai et al. (2014). The sham intervention, which involved faking the IAT, turned out to be the most effective ($d = 1.03$), followed by imagining a vivid counterstereotypic scenario ($d = 0.52$) and using implementation intentions ($d = 0.44$). However, only one intervention was effective after a delay of several hours to several days—forming implementation intentions, which had a small effect in the follow-up ($d = 0.13$, 95% CI = [0.01, 0.24]). All of the remaining interventions were ineffective at inducing lasting change in implicit bias, despite well-powered samples. Regarding explicit bias, all but one intervention (practicing the IAT with counter-stereotypical exemplars, $d = 0.15$) failed to show effects. The correlation between implicit and explicit racial bias was small ($r = .17$), thus suggesting that these variables may tap into quite different psychological processes. Lai et al. also found that the interventions did not significantly impact support for affirmative action. To summarise, this research suggests that, despite some interventions changing implicit bias in the short term, it is unlikely that most of them can produce sustained change in implicit racial preferences, and it is even less likely that they can change explicit bias or political preferences.

6.3 The Present Research

Building on the research by Devine et al. (2012), Lai et al. (2016), and Forscher et al. (2017), the present research aimed to evaluate the potential effects of counter-

stereotypical experiences (i.e., information, situations, and/or events that activate multiple, unusual, and unexpected social categories, see Section 2.2.3) on cognitive flexibility, explicit bias, concern for discrimination, and epistemic unfreezing. In Chapters 4 and 5 of this thesis, 'one-shot' CST interventions were employed in order to examine their effects on cognitive reflection (a relative of cognitive flexibility). The experiment described in this chapter, in contrast, will involve an intervention that repeatedly exposes people to counter-stereotypical information. The experiment was designed to span two weeks, which was deemed a sufficiently long period for the predicted effects to arise (similar to the 14-day intervention by Forscher et al., 2017). As such, the primary aim of this research is to test a longitudinal prediction made by the CPAG model, which is that perceivers will develop generalised cognitive flexibility if they are repeatedly exposed to counter-stereotypical diversity.

A second aim of the present research was to extend the scope of existing work by including measures of explicit bias, concern for discrimination, and a process termed epistemic unfreezing, which I define as a psychological state characterised by curiosity and open-mindedness. Regarding explicit bias, the present work capitalises on work reporting direct effects of exposure to CSTs on explicit prejudice (see Chapters 3 and 4). The goal was to replicate the finding that exposure to CSTs reduces explicit bias (i.e., negative attitudes and feelings towards outgroups). Furthermore, as described above, concern for discrimination is a motivational variable that may be negatively affected by counter-stereotypical imagery (Forscher, Mitamura, Dix, Cox & Devine, 2017)——a finding that the present research aimed to replicate as well. Finally, epistemic unfreezing is a construct that has seen some, but

relatively little, empirical attention. One line of research by Tadmor, Hong, Chao, Wiruchnipawan, and Wang (2012) tested whether multicultural experiences could decrease intergroup bias via the psychological mechanism of epistemic unfreezing. Here, epistemic unfreezing was defined as the motivation to search and process information deeply, and it was operationalised as a reduced need for cognitive closure. In the present research, epistemic unfreezing is conceptualised slightly differently by focusing on its motivational (curiosity, interest, and state openness) rather than cognitive underpinnings (the desire for cognitive closure) because the motivational consequences of counter-stereotypical experiences are underexplored. The goal was to explore whether repeated exposure to CSTs could induce epistemic unfreezing (i.e., increases in curiosity, interest, and open-mindedness) over time.

A third and final aim of the present research was to provide a conceptual replication of the finding reported in past research suggesting that Personal Need for Structure (PNS) moderates the effects of counter-stereotypical experiences on cognitive flexibility (Gocłowska & Crisp, 2013; Gocłowska et al., 2014). Taken together then, this research has a confirmatory part (aiming to test effects on cognitive flexibility, explicit bias, concern for discrimination, and the role of PNS) and an exploratory part (exploring effects on epistemic unfreezing). The next section describes the specific hypotheses that were tested.

[2]The preregistration can be found here: https://osf.io/ghkat/register/565fb3678c5e4a66b5582f67.

6.3.1 Hypotheses

The hypotheses, data collection procedures, and analysis plan were preregistered on the Open Science Framework.[2] The following directional predictions were made.

Cognitive flexibility. It was predicted that repeated exposure to CSTs would enhance cognitive flexibility over time. Specifically, a process of cognitive adaptation was expected that involves both a linear and a quadratic component. That is, early in the intervention, it was expected that participants would feel surprised by the CSTs because the CSTs will be perceived as novel. Exposure to CSTs was therefore predicted to have a relatively powerful effect on cognitive flexibility early in the intervention such that cognitive flexibility would increase. However, it was also expected that participants' surprise would fade as they get repeatedly exposed to CSTs. As a result, the effect of exposure on cognitive flexibility was expected to "flatten" towards the end of the intervention.

Explicit bias. It was predicted that repeated, longitudinal exposure to CSTs would reduce explicit bias towards different outgroups over time. Specifically, a process of cognitive adaptation that involves both a linear and a quadratic component was expected, with a similar pattern to the effects on cognitive flexibility described above.

Concern for discrimination. In line with Forscher et al. (2017), concern for discrimination was expected to decrease after exposure to CSTs because the existence of CSTs may suggest that prejudice and discrimination have become less of an issue in contemporary society. In other words, reading about individuals who challenge

stereotypes and norms may lead people to think that discrimination is not a problem that requires as much attention as in the past.

Mediation. It was predicted that the effect of exposure to CSTs on prejudice would be mediated by an increase in cognitive flexibility. This prediction is grounded in theories and research showing that counter-stereotypical experiences can prompt a switch from heuristic, category-based thinking toward more systematic, flexible thinking and thus can reduce prejudice (Crisp & Turner, 2011; Prati et al., 2015b). This prediction is also in line with the argument that rigid thinking (or need for cognitive closure) represents the motivated cognitive basis of prejudice (Roets & Van Hiel, 2011). Thus, if exposure to CSTs leads to increased cognitive flexibility (or decreased rigid thinking), then it is also likely to lead to decreased prejudice.

Moderation. It was expected that Personal Need for Structure (PNS) would moderate the effect of condition on cognitive flexibility and prejudice, respectively. More precisely, participants low in PNS were expected to show the greatest increases in cognitive flexibility, whereas participants high in PNS were expected to show the reverse effect, that is, decreases in cognitive flexibility (Gocłowska & Crisp, 2013; Gocłowska et al., 2014).

6.4 Method

6.4.1 Pilot Experiment

An exploratory pilot study pretested 30 counter-stereotype manipulations, which were newly constructed by the author and inspired by real life counter-stereotypical individuals (e.g., Muslim feminists, atheists who study Muslim religion). A pilot study determined the ten most effective manipulations to use in Experiment 12 (similar to the method used in the pilot studies described in Chapter 4). The required sample size was determined based on the same criteria as the pilot studies in Chapter 4. A total of 266 participants (39% male, 60% female, 1% other; M_{age} = 38.46, SD_{age} = 11.69) was recruited via Prolific (www.prolific.ac) to take part in a study "exploring how people form first impressions of other people based on limited information". They were assigned to one of six conditions (gender: Typical vs. atypical target individuals, political ideology: Typical vs. atypical target individuals, and Muslim religion: Typical vs. atypical target individuals) and in each condition they were presented with information about ten target individuals. The manipulations were constructed such that the content was identical across conditions, and only the category memberships were crossed (see Table 6.1 for examples). This structure of the manipulation allowed for a rigorous test of the effects of counter-stereotypical experiences, because it could help to rule out the possibility that differences in terms of content produced any effects.

Following the description of each individual, participants were asked "To what

Table 6.1: Between Day 2 and Day 10, participants were randomly presented with two of the ten exemplars (e.g., stereotypes 1+2 in the typical condition and counter-stereotypes 1+2 in the atypical condition). The colour scheme highlights how the content always stayed identical across conditions——conditions only varied in terms of how the categories were crossed.

Stereo-typical condition	**Stereotype 1:** "Mary works as a midwife, and she loves shopping with her friends."	**Stereotype 3:** "Zainab wears a headscarf, and she reads the Quran every day."	**Stereotype 5:** "Rebecca wants to do away with tuition fees for students, and she thinks that immigration and diversity are great."	**Stereotype 7:** "Ahmed believes that there is only one God, Allah, and he fasts once a year for a month from dawn until sunset."	**Stereotype 9:** "Helen thinks that believing in life after death is absurd, and that accepting religious beliefs uncritically is harmful."
	Stereotype 2: "Paul works as a car mechanic, and he loves playing football."	**Stereotype 4:** "John doesn't believe in God, and he thinks that women should be able to freely choose abortion."	**Stereotype 6:** "Patrick thinks that businesses need less regulation, and he thinks trade unions have too much power."	**Stereotype 8:** "Lisa doesn't follow any religion, and she wants more equality for women."	**Stereotype 10:** "Abdul believes in the prophets of Allah, and he performs five formal prayers a day in a Mosque."
Counter-stereo-typical condition	**Counter-stereotype 1:** "Mary works as a car mechanic,and she loves shopping with her friends."	**Counter-stereotype 3:** "Zainab wears a headscarf,and she thinks that women should be able to freely choose abortion."	**Counter-stereotype 5:** "Rebecca wants to do away with tuition fees for students,and she thinks that businesses need less regulation."	**Counter-stereotype 7:** "Ahmed believes that there is only one God, Allah, and he wants more equality for women."	**Counter-stereotype 9:** "Helen thinks that believing in life after death is absurd,and she performs five formal prayers a day in a Mosque."
	Counter-stereotype 2: "Paul works as a midwife,and he loves playing football."	**Counter-stereotype 4:** "John doesn't believe in God,and he reads the Quran every day."	**Counter-stereotype 6:** "Patrick thinks that immigration and diversity are great,and he thinks trade unions have too much power."	**Counter-stereotype 8:** "Lisa doesn't follow any religion,and she fasts once a year for a month from dawn until sunset."	**Counter-stereotype 10:** "Abdul believes in the prophets of Allah,and that accepting religious beliefs uncritically is harmful."

extent do you feel surprised?" and were asked to indicate their response on a scale ranging from 1 (*not at all*) to 7 (*very much*). It was predicted that the atypical target individuals (i.e., the female car mechanic, the liberal Muslim, etc.) would be rated as significantly more surprising than typical target individuals (i.e., the male car mechanic, the traditional Muslim, etc.). In support of this idea, there was a statistically significant effect of condition on ratings of surprise, $t(251) = 7.86$, $p(\text{one-tailed}) < .0005$, $d = 0.96$, which is a large effect. As expected, participants in the atypical condition rated the target individuals as significantly more surprising ($M = 3.38$, $SD = 1.12$) than participants in the typical condition ($M = 2.34$, $SD = 1.01$).

6.4.2 Experiment 12

Participants. The required sample size for the main study was determined based on Fritz and MacKinnon's (2007) article, which calculates minimum required sample sizes for mediation models depending on the size(s) of the hypothesised effects. Assuming small-to-medium sized α and β paths in the predicted mediation model (i.e., $\alpha = .26$ and $\beta = .26$) and 80% power, I aimed to recruit a minimum of $N = 148$ participants. A total of 197 British participants (32% male, 68% female; $M_{age} = 35.82$, $SD_{age} = 10.67$) were recruited via Prolific. The dropout rate was relatively low (22%), meaning that, on average 78% of participants completed any given session of the study. Each participant was expected to spend approximately 45 minutes completing all parts of the study. More precisely, parts 1 and 7 took about 5 minutes

to complete, respectively, whereas parts 2 through 6 took about 7 minutes to complete, respectively. To ensure that participants completed each session in due course (i.e., to minimise the chances of them getting distracted by other tasks), maximum allowed completion times were specified via Prolific (15 minutes for the baseline and follow-up sessions and 17 minutes for the intervention sessions, respectively). Participants were paid £4.40 each in total for their time. In addition, by promising to pay a bonus of £0.50 to every participant who successfully completed all parts of the study, participants had an incentive to complete all experimental sessions.

Participant inclusion criteria were specified a-priori as follows: (1) Participants' country of birth = United Kingdom, (2) Participants' nationality = British, (3) Participants' current country of residence = United Kingdom, (4) Participants' age = at least 18 years old, and (5) participants had to have completed at least 1 study on Prolific (to ensure that the participants were familiar with Prolific's system and knew how to navigate it). In addition, participants were excluded if (6) they had previously participated in any of the studies run by the author via Prolific, (7) they did not take the study seriously, that is, if they gave empty or nonsensical responses, then they were excluded from the respective time point, (8) failed the attention check (these participants were excluded from the respective time point, but not necessarily from the study as a whole), (9) guessed the purpose of the study (in order to minimise demand characteristics),

Procedure. Over the course of 14 days, participants were invited to complete an online questionnaire every two days. On Day 0, baseline measures of cognitive flexibility, explicit bias, concern for discrimination, PNS, and demographic vari-

ables (gender, age, nationality) were taken. On Day 2, participants were randomly assigned to the experimental versus control condition. Participants in the experimental condition were presented with information about different individuals who challenged stereotypes, whereas participants in the control condition were presented with information about different individuals who fitted stereotypes. On Days 4 to 10 participants remained in their respective conditions and were presented with further information that either challenged or fitted stereotypes. Note that measures of all dependent variables were answered on each occasion, that is, not only pre- and post-intervention. On Day 12, the final time point, participants were not presented with any information about individuals, but completed follow-up measures of cognitive flexibility, explicit bias, and concern for discrimination.

At the beginning of each experimental session, participants were asked to indicate their feelings towards different social groups using feeling thermometers, which are a validated measure of explicit bias (e.g., Forscher et al., 2017; Meleady, Seger, & Vermue, 2017). Participants read: "We would like to know how 'cold' (unfavourable) or 'warm' (favourable) you feel towards different social groups. Please indicate your feelings towards each social group listed below using the scale provided." Participants then rated nine different social groups (the elderly, the disabled, Catholics, Muslims, atheists, HIV patients, asylum seekers, gay men, and EU immigrants) on a scale from 0 (*very cold*) to 100 (*very warm*). Only three of the social groups were directly relevant to the experiment (Muslims, atheists, and asylum seekers); all remaining groups represented filler items. Note that six out of the ten manipulations crossed characteristics pertaining to Muslims and atheists, so it was expected that

the predicted effects would be most pronounced for these two outgroups. Asylum seekers are not a group that is directly related to the manipulations, but it was expected that the predicted effects would generalise to this outgroup, in line with previous research suggesting that exposure to CSTs can promote generalised tolerance that extends towards unrelated outgroups (Vasiljevic & Crisp, 2013).

On Day 2, after being randomly assigned to the experimental versus control condition, participants were asked to vividly imagine the life and experiences of two target individuals (from the set presented in Table 6.1) and to describe what they imagined. Samples of imagery descriptions that were produced by participants can be found in Appendix F. The information about the target individuals was always presented in a random order and in a counterbalanced fashion. Following the imagination task, participants answered two items, that is, "To what extent do you feel surprised?" and "To what extent is [name] unusual?" on a scale from 1 (*not at all*) to 7 (*very much*), which were later combined to form the manipulation check ($\alpha = .89$ across all time points). The participants then answered three items measuring epistemic unfreezing ("I would like to learn more about [name]"——reflecting interest, "I'm feeling curious about what [name] is like"——reflecting curiosity, and "I'm feeling inquisitive, open-minded"——reflecting openness), also on a 7-point scale ($\alpha =$

[3]Internal consistency was also calculated for the manipulation check items at each time point, which rendered the following results: α(Day 2) = .86, α(Day 4) = .89, α(Day 6) = .87, α(Day 8) = .91, α(Day 10) = .93. Similarly, internal consistency was calculated for the epistemic unfreezing items at each time point, which rendered the following results: α(Day 2) = .86, α(Day 4) = .88, α(Day 6) = .90, α(Day 8) = .93, α(Day 10) = .94.

.90 across all time points).[3]

Next, in order to measure cognitive flexibility, participants completed a task from Gocłowska and Crisp (2013) and Boot et al. (2017). Specifically, participants were given two minutes to come up with as many names as possible for a new kind of product, venture, or item. The instructions read: "We would like you to think of various names for a new kind of PASTA, for example 'fussilini', 'paragoni', 'malwini'. What other names can you come up with? In the box below, please write down ALL the names that come to your mind." At each time point (i.e., Day 0, Day 2, Day 4, Day 6, Day, 8, Day 10, and Day 12), participants were randomly presented with a different version of the task (one per time point). In addition to names for pasta, they were asked to come up with names for planets (e.g., verunus, arctanus, tronus), flowers (e.g., lunia, fridia, ezilia), software companies (e.g., Triddle, Wubble, Kimple), airlines (e.g., Jimair, Greenair, Scanair), Greek islands (e.g., Mianos, Nikonos, Presos), and cocktails (e.g., domicita, hawaiana, passilada). Participants' responses to the naming tasks were coded in order to determine their levels of cognitive flexibility. A coder, blind to condition, employed the so-called *snapshot scoring technique* (Silvia, Martin & Nusbaum, 2009) that involved scoring participants' responses in terms of how diverse and dissimilar they were on a scale from 1 (*not at all diverse*) to 5 (*very diverse*). The detailed scoring instructions can be found in Appendix G. A second coder, also blind to condition, coded another 10% of the material in order to determine the degree of inter-rater reliability. The results revealed an inter-rater reliability of $\alpha = .68$, which was slightly below the required threshold ($\alpha = .7$) specified in the preregistration. It became

apparent that the two coders had used slightly different benchmarks when scoring participants' responses—the scores given by the second coder were consistently more positive/optimistic than those by the first coder. To resolve this discrepancy the first coder corrected all scores by adjusting them upwards where she saw fit, always blind to condition. To check whether the adjusted scores were more appropriate than in the first round of coding, a third coder then rated a different 10% of responses. The new interrater reliability coefficient (computed for the first and third coder) was $\alpha = .73$, which was sufficiently high. The adjusted scores given by the first coder were used to measure cognitive flexibility.

Finally, participants responded to eight items measuring concern for discrimination on a scale from 1 (Strongly disagree) to 10 (Strongly agree; adapted from Forscher et al., 2017). Four of the items measured religious discrimination (items of interest), whereas the remaining four items measured gender discrimination (filler items). Example items included "People make more fuss about discrimination against Muslims than is necessary" and "I consider religious discrimination to be a serious social problem".

On Day 0 (baseline) participants completed the PNS scale (Neuberg & Newsom, 1993), answered some demographic questions (sex, age, nationality), and were given the opportunity to provide feedback on the experiment (e.g., report any technical difficulties). An index reflecting PNS was computed by averaging the 12 PNS items (reverse-coded where appropriate, $\alpha = .95$). On Day 12 (follow-up) participants were asked about their thoughts on the purpose of the study, thanked for participating, and debriefed.

Note that the measure of prejudice was always presented at the beginning of the sessions, prior to the manipulation, and cognitive flexibility was always measured after the manipulation. This design ensured that prejudice could only be affected by the previous session's manipulation and the previous session's level of cognitive flexibility, thus providing a stronger test of the predicted mediation effect. Furthermore, an attention check was included at each time point, as part of the scale measuring concern for discrimination: "It is important that you pay attention to this study. Please tick 'Strongly agree'."

To ensure that participants completed the questionnaires at relatively even time intervals, they were instructed to complete them between 12pm and 8pm British time. In each session, one reminder was sent halfway through the time window (i.e., at 4pm) to those participants who had not yet completed that day's session. In addition, participants were instructed not to complete the study using a mobile device, but from a laptop or desktop computer instead.

6.4.3 Data Analytic Plan

Analyses were conducted using mixed effects models with the lme4 package in *R* (Bates, Mächler, Bolker & Walker, 2015) and graphs were plotted using ggplot2 (Wickham, 2009). Linear mixed effects models were used throughout because all of the dependent variables were quantitative. To estimate the effects of the intervention on the dependent variables, confidence intervals were obtained using likelihood profiles. Because mixed effects models use all of the information that is provided

by each participant, they provide a natural framework for handling missing data as long this data is missing at random (Ibrahim & Molenberghs, 2009). The random effects structure always included a random intercept for each participant, and random slopes were always included for independent variables that varied within participants. The independent variables of all models included the linear and quadratic effects of time and indicators for condition and the condition by time interactions, and variables were centred prior to computing the interactions. When time was included as a independent variable, the models also always contained random slopes for time. Time was scaled such that each unit represented one day. Therefore, in all of the analyses of the overall effects of the intervention, the linear effect of time represents the degree to which the outcome of interest changed each day, the quadratic effect of time represents the degree to which the rate of change over time accelerated or decelerated each day, and the interactions between the time contrasts and condition represent the difference between the intervention and control groups in their rate of change and acceleration/deceleration per day.

The analyses had five goals: To provide a conceptual replication of (1) the direct effect of exposure to CSTs on cognitive flexibility reported in past research (Di Bella & Crisp, 2016; Gocłowska et al., 2012; Vasiljevic & Crisp, 2013; Prati et al., 2015b) and (2) the finding the PNS moderates this effect (Gocłowska & Crisp, 2013; Gocłowska et al., 2014), (3) the finding that exposure to CSTs can reduce prejudice/explicit bias (Hall & Crisp, 2005; Vasiljevic & Crisp, 2013), (4) the finding that exposure to CSTs can reduce concern for discrimination (Forscher et al., 2017), and also to conduct exploratory analyses examining (5) whether exposure to CSTs

affects epistemic unfreezing.

6.5 Results

6.5.1 Effects at Baseline

Participants in the experimental and control conditions did not differ in any of the dependent variables at baseline (see Table 6.2), so it seems that random assignment to conditions was successful. However, more participants completed sessions in the control than in the experimental condition (b = 0.43, $t(196)$ = -1.97, p(two-tailed) = .05. This result may be due to chance or it may represent a systematic difference between conditions—at present, it is difficult to interpret what caused the discrepancy. In Section 6.6.2, this result and its implications will be further discussed.

6.5.2 Manipulation Check

First, the effectiveness of the manipulation was tested. The mean of the scores of the two manipulation check items ("To what extent do you feel surprised?" and "To what extent is [name] unusual?" on a scale from 1 (*not at all*) to 7 (*very much*)) were combined into one index called "Surprise". The results revealed a large effect of the manipulation: Participants differed significantly in terms of how surprising and unusual they thought the counter-stereotypical individuals to be, b = 1.457, 95% CI = [1.157, 1.756]. There were no other statistically significant effects.

Table 6.2: Characteristics of the intervention and control groups at baseline.

	Control				Intervention				Difference
Rating	***N***	***Mean***	***SD***	***Skew***	***N***	***Mean***	***SD***	***Skew***	
Flexibility	96	3.35	1.26	-0.28	98	3.56	1.31	-0.56	0.207 [-0.158, 0.572]
Muslims	97	56.19	27.45	0.02	101	56.35	26.31	-0.26	0.161 [-7.374, 7.696]
Atheists	97	68.08	24.93	-0.39	101	70.42	21.59	-0.20	2.333 [-4.195, 8.861]
Asylum seekers	97	52.86	27.82	-0.07	101	52.90	29.19	-0.10	0.045 [-7.954, 8.044]
Concern	97	5.86	2.13	-0.07	101	6.41	2.08	-0.36	0.550 [-0.040, 1.140]
PNS	97	3.80	0.74	0.10	100	3.93	0.71	0.16	0.128 [-0.075, 0.332]

Note: The "Difference" column represents the mean difference in rating and its 95% CI. "Muslims" signifies prejudice against Muslims, "Atheists" signifies prejudice against atheists, and "Asylum seekers" signifies prejudice against asylum seekers.

6.5.3 Confirmatory Analyses: Effects on Cognitive Flexibility, Explicit bias, and Concern for Discrimination

Participants in both conditions showed a decrease in cognitive flexibility of 0.027 units a day on average, 95% CI = [-0.042, -0.011]. This effect did not differ across conditions, b = -0.243, 95% CI = [-0.529, 0.044]. Similarly, no statistically significant effect of condition was found for explicit bias against Muslims and atheists, although a linear effect of time on explicit bias against asylum seekers was detected such that, on average, participants in both conditions showed increased feelings of warmth by 0.196 units a day, 95% CI = [0.024, 0.369]. No statistically significant effects of condition or time on concern for discrimination were found. The results are shown

Table 6.3: Change over time in surprise, cognitive flexibility, prejudice against Muslims, atheists, and asylum seekers, and concern for religious discrimination.

Rating	**Condition**	**Time**	**Time squared**	**Condition by time**	**Condition by time squared**
Surprise	1.479 [1.178, 1.780]*	-0.022 [-0.067, 0.023]	0.000 [-0.015, 0.014]	-0.026 [-0.118, 0.063]	0.004 [-0.026, 0.034]
Flexibility	-0.243 [-0.529, 0.044]	-0.027 [-0.042, -0.011]*	-0.002 [-0.006, 0.002]	-0.027 [-0.053, 0.008]	-0.002 [-0.001 0.016]
Muslims	-0.25 [-8.203, 7.698]	-0.099 [-0.299, 0.102]	-0.005 [-0.038, 0.027]	-0.065 [-0.465, 0.336]	-0.002 [-0.067, 0.063]
Atheists	-0.405 [-6.574, 5.768]	-0.104 [-0.283, 0.075]	0.008 [-0.031, 0.047]	-0.098 [-0.456, 0.259]	0.059 [-0.018, 0.136]
Asylum seekers	-2.224 [-10.530, 6.087]	0.196 [0.024, 0.369]*	-0.031 [-0.062, 0.001]	0.01 [-0.335, 0.355]	0.056 [-0.007, 0.119]
Concern	0.224 [-0.404, 0.852]	-0.006 [-0.027, 0.016]	-0.001 [-0.004, 0.003]	-0.02 [-0.063, 0.023]	0.005 [-0.003, 0.013]

Note: Estimates and their profile likelihood 95% CIs were derived from Linear Mixed Effects Models (LMEMs) containing condition, linear time, quadratic time, and the interactions between condition and linear/quadratic time. All models contained a random intercept for each participant and random slopes for time and quadratic time. "Muslims" signifies prejudice against Muslims, "Atheists" signifies prejudice against atheists, and "Asylum seekers" signifies prejudice against asylum seekers. * signifies statistical significance $< .05$.

in Figure 6.1 and Table 6.3.

These results do not support the predictions made regarding cognitive flexibility, explicit bias, and concern for discrimination, despite the manipulation check confirming that the manipulation was effective. Given this lack of direct effects, testing the predicted mediation effect was deemed redundant.

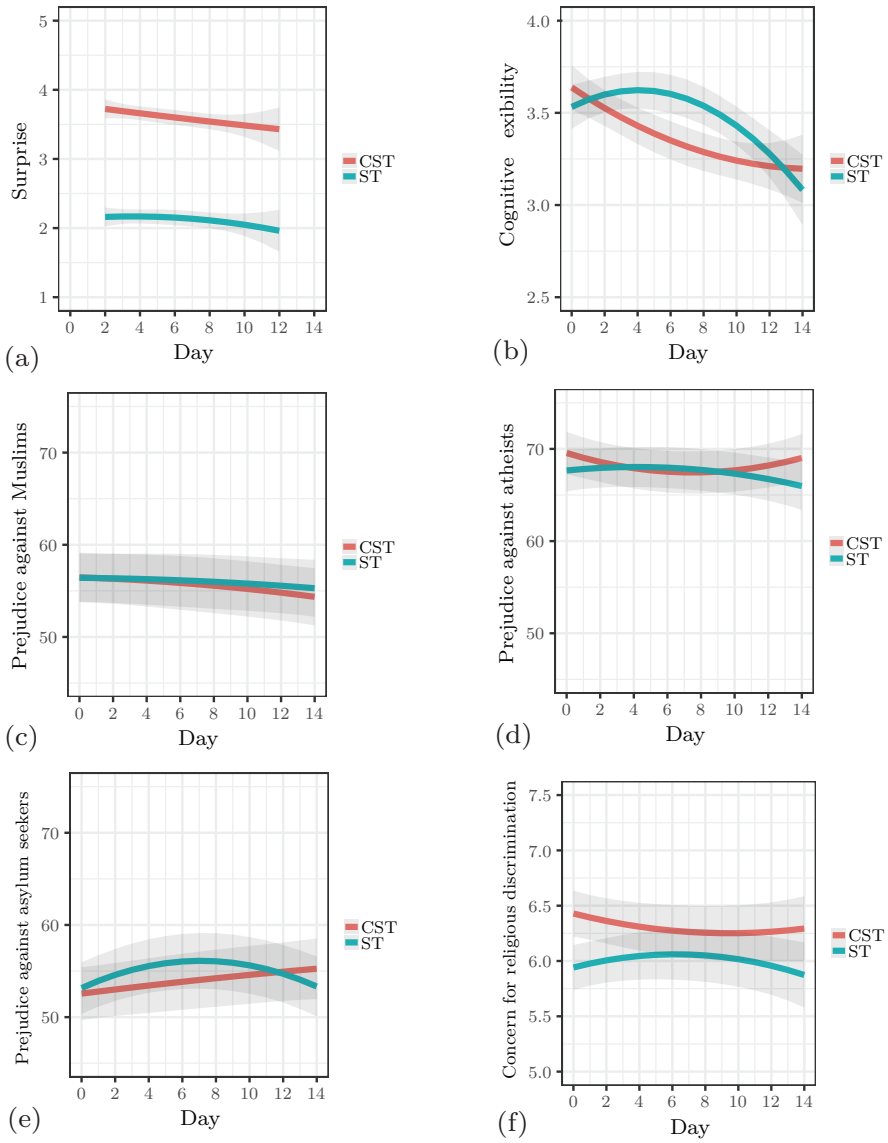

Figure 6.1: The effects of the intervention (CST = counter-stereotypical or intervention condition; ST = stereotypical or control condition) on (a) surprise, (b) cognitive flexibility, (c) prejudice against Muslims, (d) prejudice against atheists, (e) prejudice against asylum seekers, and (f) concern for discrimination, depicted over time. Day 0 represents the baseline and day 2 represents the start of the intervention, which was carried out until day 10. On Day 12, follow-up measures of the dependent variables were taken.

6.5.4 Aggregate Confirmatory Analyses: The Moderating Role of PNS

To test whether the intervention had an effect on cognitive flexibility across all time points, and whether the hypothesised direct effect was moderated by PNS, moderated regression analyses were computed using the pequod package in *R* (Mirisola & Seta, 2016). Conditions were contrast coded as –1 (control) and +1 (experimental). To test for the effect of the intervention on cognitive flexibility, I regressed the cognitive flexibility index on condition, PNS, and their interaction. The independent variables were centred prior to computing the interaction term.

First, the analyses revealed a main effect of PNS on cognitive flexibility (b = -0.20, $t(193)$ = -2.31, p = .02), such that participants low in PNS showed higher cognitive flexibility (M = 3.90, SD = .62) than participants high in PNS (M = 3.40, SD = .77). Further, there was no effect of condition on cognitive flexibility (b = -0.56, $t(193)$ = -1.63, p = .11). Whereas participants in the experimental condition had an average flexibility rating of 3.41 (SD = .91), participants in the control condition had an average flexibility rating of 3.51 (SD = .89). This pattern of results does not support the prediction that counter-stereotypical experiences boost cognitive flexibility.

In order to test whether the intervention had a different effect on participants low versus high in PNS——such that participants low in PNS would show increased cognitive flexibility, whereas participants high in PNS would show decreased cognitive flexibility——an interaction term was computed between condition and PNS,

with condition and PNS as covariates, respectively. To understand the nature of the interaction, the effect of condition on cognitive flexibility at different levels of PNS was inspected using simple slopes analysis (Aiken & West, 1991). "Low PNS" was defined as 1SD below the mean and "high PNS" as 1SD above the mean. The results revealed a non-significant interaction effect (b = 0.14, $t(193)$ = 1.55, p = .12). Similarly, simple slopes analyses did not show any significant effects. Neither participants low in PNS (b = -0.14, 193) = -1.50, p = .14) nor participants high in PNS (b = 0.06, $t(193)$ = .69, p = .49) seemed to perform differently across conditions. Taken together, the above results do not provide support for the prediction that PNS moderates the effects of exposure to CSTs on cognitive flexibility.

6.5.5 Exploratory Analyses: Effects on Epistemic Unfreezing

Exploratory analyses were computed to test whether the intervention had an effect on epistemic unfreezing, which was operationalised by creating an index of the items reflecting curiosity, interest, and openness. The analyses revealed an effect of condition on epistemic unfreezing, b = 0.382, 95% CI = [0.003, 0.760]. In addition, the results showed that epistemic unfreezing linearly declined in both conditions over time, b = -0.077, 95% CI = [-0.103, -0.051]. No other significant effects were found. This suggests that participants in the experimental condition were, on average, more "epistemically unfrozen" than were participants in the control condition. However, these results also suggest that participants in both conditions gradually

became less epistemically unfrozen over time, which is in line with the prediction that repeatedly imagining target individuals would be experienced as decreasingly novel and interesting.

To explore the effect of condition on curiosity, interest, and openness, further analyses were computed with each of the items as dependent variables. The analyses revealed that the intervention had an effect on interest (b = 0.437, 95% CI = [0.022, 0.850]) and curiosity (b = 0.451, 95% CI = [0.033, 0.869]), but not on openness (b = 0.266, 95% CI = [-0.150, 0.680]). The results are shown in Table 6.4 and Figure 6.2. In sum, it appears that participants in the intervention (vs. control) group felt more curious about and interested in the target individuals, but were not necessarily more open-minded and inquisitive as a result of being exposed to CSTs. In addition, there was a linear effect of time such that participants in both conditions showed decreases in epistemic unfreezing and in all of its hypothesised component processes.

6.6 General Discussion

The empirical work described in this chapter examined whether repeated exposure to CSTs can enhance cognitive flexibility, reduce explicit bias, and reduce concern for discrimination over time. The longitudinal experiment presented here did not replicate many of the effects reported in the literature (Blair et al., 2001; Gocłowska et al., 2012; Kunda & Oleson, 1995; Prati et al., 2015a; Vasiljevic & Crisp, 2013). That is, the confirmatory analyses showed no statistically or practically significant effects of the intervention on cognitive flexibility, explicit bias, and concern

Table 6.4: Change over time in epistemic unfreezing and its hypothesised component processes (interest, curiosity, and openness).

Outcome	Condition	Time	Time squared	Condition by time	Condition by time squared
Unfreezing	0.382 [0.003, 0.760]*	-0.077 [-0.103, -0.051]*	0.008 [-0.002, 0.018]	0.034 [-0.018, 0.086]	-0.008 [-0.028, 0.011]
Interest	0.437 [0.022, 0.850]*	-0.068 [-0.098, -0.037]*	0.006 [-0.005, 0.018]	0.053 [-0.007, 0.114]	-0.011 [-0.033 0.012]
Curiosity	0.451 [0.033, 0.869]*	-0.071 [-0.101, -0.040]*	0.007 [-0.005, 0.018]	0.044 [-0.016, 0.105]	-0.008 [-0.032, 0.015]
Openness	0.266 [-0.150, 0.680]	-0.093 [-0.124, -0.063]*	0.012 [0.000, 0.023]	0.012 [-0.050, 0.074]	-0.004 - [0.026, 0.018]

Note: Estimates and their profile likelihood 95% CIs were derived from Linear Mixed Effects Models (LMEMs) containing condition, linear time, quadratic time, and the interactions between condition and linear/quadratic time. All models contained a random intercept for each participant and random slopes for time and quadratic time.
* signifies statistical significance $< .05$.

for discrimination despite strong experimental manipulations, which produced the expected changes in surprise in line with past research showing that exposure to CSTs causes surprise (Hutter, Wood & Dodd, 2012; Prati et al., 2015a). Similarly, no effects were found for PNS as a moderating variable (Gocłowska & Crisp, 2013; Gocłowska et al., 2014) and the intervention did not seem change participants' level of concern about discrimination (Forscher et al., 2017).

Intriguingly, clear effects were found for the exploratory analyses examining the effects of the intervention on epistemic unfreezing: Exposure to CSTs made participants more interested in and curious about the target individuals, but this effect did not extend beyond the immediate context to more general openness. That is, effects were found for items that referred to the target individuals directly ("I'm

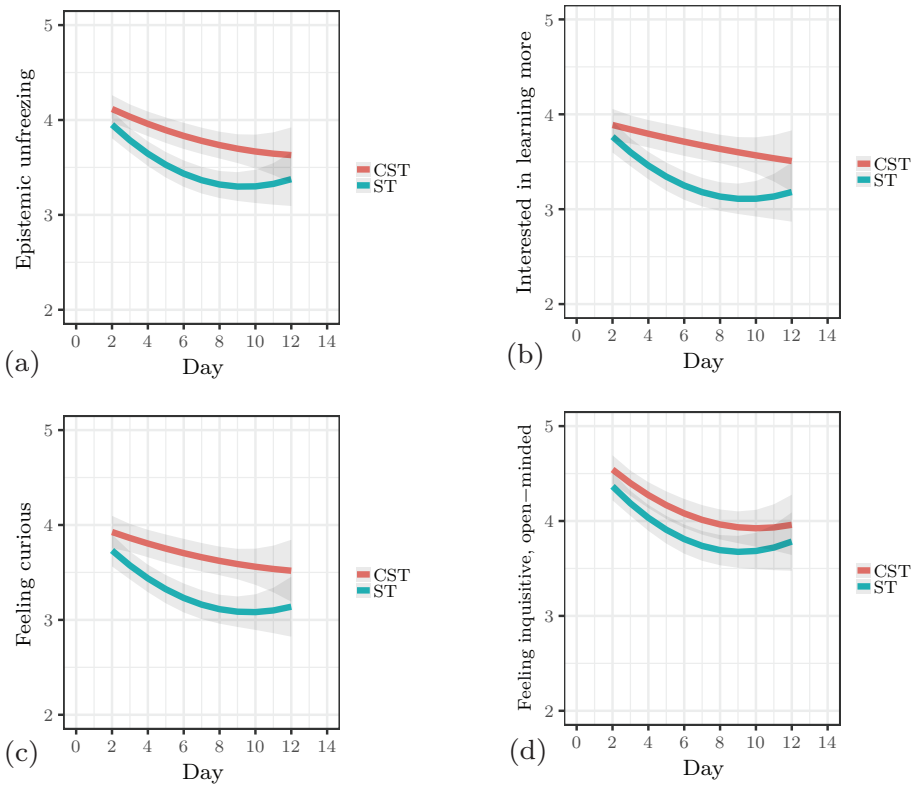

Figure 6.2: The effects of the intervention (CST = counter-stereotypical or intervention condition; ST = stereotypical or control condition) on (a) epistemic unfreezing, (b) interest, (c) curiosity, and (d) openness, depicted over time.

interested in learning more about [name]", "I'm feeling curious about [name]"), but not for items that were unrelated to the situation at hand ("I'm feeling inquisitive, open-minded", "I consider religious discrimination to be a serious social problem"). Thus, there appears to be no cognitive transfer, or generalisation, beyond the immediate situation from the individual to the group level, which suggests that the feelings and attitudes towards target individuals may not easily generalise towards different outgroups that the target individuals are affiliated with. On the one hand,

this is not too surprising because perceivers often treat counter-stereotypical individuals as exceptions to the rule and so "subtype" them (i.e., do not view them as representative of the target category; Johnston & Hewstone, 1992, Kunda & Oleson, 1995). On the other hand, previous research suggests that exposure to CSTs can promote generalised tolerance and egalitarianism that extends beyond the immediate counter-stereotypical situation (Prati et al., 2015a; Vasiljevic & Crisp, 2013). A tension exists between these two perspectives, which only future research may be able to resolve by, for example, measuring subtyping and exploring its role in the hypothesised effects. At present, it is difficult to evaluate whether the CST intervention employed here failed to generalise towards relevant outgroups due to subtyping, or for other reasons.

The findings of the experiment presented in this chapter echo the results reported in the preceding two empirical chapters. That is, a direct effect of counter-stereotype interventions on cognitive reflection / flexibility could not be established in the empirical work described in both Chapters 4 and 5 either. In Chapter 4, the sample-weighted average effect size across eight experiments (two in the laboratory and six online) was d_+ = -0.13, which resembles the negative trend seen in this experiment. In Chapter 5, the effect of the experimental manipulation on cognitive reflection was very small and non-significant, $d_+ = 0.08$. Unfortunately, it is not possible to draw clear conclusions from non-significant findings. Thus, further research is needed that attempts to directly (by re-running this experiment once again using a British participant sample) and conceptually (e.g., by using alternative measures of cognitive flexibility and explicit bias) replicate the reported results in order to

confirm their validity and reliability.

6.6.1 Theoretical and Practical Implications

What could explain the lack of an effect of the intervention on cognitive flexibility? First, it may be that the large effect of CST interventions on cognitive flexibility reported in past research (Gocłowska et al., 2012; Vasiljevic & Crisp, 2013; Prati et al., 2015b) is more constrained than previously thought. For example, it may be that the effect only arises only under certain conditions (e.g., when the target individuals bear direct relevance to the participants), or it may be that certain circumstances easily undermine the effect (e.g., fatigue, boredom). If so, then this has implications for models and theories that seek to understand the effects of counter-stereotypical experiences. As discussed in Chapter 4, in order to improve existing models, it may be crucial to detail the optimising and boundary conditions of the effects of counter-stereotypical experiences, and to specify the respective psychological mechanisms reducing or boosting different dependent variables. Ideally, such research would be: (1) cumulative (i.e., build directly on existing research), (2) theory-driven (for example, see Chapter 7, which introduces a new theory specifying the psychological effects of expectancy violations), (3) highly powered (to allow a precise estimation of the hypothesised effects), (4) rule out that measures of potential moderators (e.g., PNS) confound other measures in the experiment (e.g., epistemic unfreezing) or vice versa (which can be achieved by administering measures of moderators vs. state measures at different time points), (5) directly

manipulate hypothesised mediating variables (to test the directionality of effects), and (6) longitudinal in nature (to allow assessing whether effects persist over time).

Second, it is possible that the hypothesised effect of exposure to CSTs on cognitive flexibility is much smaller than previously assumed. If that is the case, then this means that (1) the present experiment may have been under-powered and (2) much larger sample sizes will be required in future research to detect the effect, if it exists. For example, if the direct effect of exposure to CSTs on cognitive flexibility were small, one would require a total of $N = 278$ to detect it (assuming 80% power and an alpha error probability $\alpha = .05$).

Third, it may be that participants felt too detached from the target individuals in the sense that they were not personally relevant to them. Thus, they may not have sufficiently engaged with the materials and only completed the imagination task superficially. While this is possible, it does not explain why past research has found medium-to-large effects of comparable CST interventions on cognitive flexibility (Gocłowska et al., 2012; Vasiljevic & Crisp, 2013; Prati et al., 2015b).

Whereas none of the confirmatory analyses yielded the predicted effects, the exploratory analyses did reveal some interesting effects. First, exposure to CSTs not only increased surprise, which is in line with past research (Hutter et al., 2012; Prati et al., 2015a), but it also helped to boost participants' interest in and curiosity towards the target individuals. The latter finding is consistent with early research, which has implied but not directly measured whether or not counter-stereotypical information could promote curiosity (Hastie et al., 1990; Kunda, 1990). This finding——i.e., that exposure to CSTs can promote interest and curiosity——is particu-

larly promising because the motivational consequences of counter-stereotypical experiences, and intergroup contact more generally, are relatively under-explored. If this result can be replicated, then it might open up a new area of research on the consequences of diversity for curiosity. Curiosity can be a very powerful driver of learning, motivation, and development (Litman, Crowson & Kolinski, 2010; Silvia, 2008; von Stumm et al., 2011), so it may be that researchers could capitalise on these observations in order to improve intergroup relations and successfully manoeuvre the challenges of rising diversity.

On a practical level, the counter-stereotype paradigm developed in the empirical work described in this chapter can be used in future research to test further predictions regarding the effects of counter-stereotypical experiences (and to retest the current predictions). By ensuring that differences in content cannot explain any observed effects, this paradigm neatly manipulates exposure to CSTs. However, the findings reported in this chapter raise an important caveat: More research, both in the laboratory and in the field, needs to be done before CST interventions can be applied in real life settings and confident predictions made about their likely effects, such as decreased intergroup bias. In line with Chapters 4 and 5, it appears that implementing CST interventions in real life settings, such as schools or companies, may be premature at this stage.

6.6.2 Limitations and Future Directions

In addition to conceptual issues (as detailed above), there are also methodological reasons that could explain the failure to replicate past findings. First, it is possible that some of the experimental materials lacked construct and/or content validity. For example, despite the fact that the naming tasks were attempts to operationalise what is thought to be the conceptual core of previous measures of cognitive flexibility (e.g., Gocłowska & Crisp, 2013), is it possible that they may not have captured the relevant construct, cognitive flexibility, adequately. Future research is required to explore the convergent and discriminant validity of the cognitive flexibility measures employed.

Second, it may be that presenting exclusively stereotype-fitting information in the control condition might have had a stereotype-enhancing effect. This is because most members of most groups do not typically comprise exclusively stereotype-confirming information. Future research should consider including a neutral control condition devoid of stereotypical information in order to examine whether the stereotypical condition inadvertently reinforces or perhaps even creates stereotypes. Such an approach would help determine whether it is stereotypicality or counter-stereotypicality driving any effects; plus, it would make the materials more representative of the realities of counter-stereotypical diversity, which is often fluid and complex.

It is also possible that some of the conditions postulated by the CPAG model were not met, which resulted in a failure to replicate the predicted effects. For

example, whether or not participants actually engaged in inconsistency resolution (i.e., stereotype suppression and generative thought; see Section 2.4.2) is unknown. Future research using similar counter-stereotype paradigms is advised to measure this process (e.g., by content-coding the imagery descriptions and/or using text mining or linguistic analysis) in order to determine when or why it fails to happen.

Why was no support found for moderating role of PNS in the effect of exposure to CSTs on cognitive flexibility, which has been reported in prior research (Gocłowska & Crisp, 2013; Gocłowska et al., 2014)? One possibility is that multiple situational and individual difference variables interact at any given moment, so taking into account only one of them, such as PNS, may not adequately capture the complexity of the psychological effects of exposure to counter-stereotypical diversity. Another possibility is that, because the present findings are surprising and unexpected, a methodological or statistical error could have caused a spurious pattern of results in this study (e.g., an incorrect calculation of the PNS index or any of the other measures). This explanation seems unlikely given the negative association between PNS and flexibility, which represents a conceptual replication of past research (e.g., Thompson et al., 1994). As such, this latter result, coupled with the successful manipulation check, provides a positive control (or, "sanity check") and suggests that the experiment likely provided a fair test of the hypotheses.

It is interesting that there seems to have been a higher rate of dropout in the intervention than in the control group. This result could be due to chance, or it could reflect a more systematic difference between the experimental conditions. Given that participants in the intervention group showed more interest in and curi-

osity towards the target individuals, they could have become pre-occupied with the imagery task and potentially even depleted by it, leading them to drop out. Previous research suggests that counter-stereotypical experiences can sometimes be experienced as cognitively effortful because the process of inconsistency resolution draws on executive functioning processes (Hutter & Crisp, 2006; Hutter et al., 2012; Macrae et al., 1999). Therefore, participants in the intervention group may have potentially been more cognitively depleted than participants in the control group. In addition, the set maximum completion times described in Section 6.4.2 ("ticking timers") may have added a sense of urgency or stress, making participants in the experimental condition feel that there was not sufficient time to make sense of the counter-stereotypical target individuals. Yet another possible explanation is that participants in the experimental condition did not like reading about CSTs (cf. research on the deviant bias in Section 2.6.1.2), or found it unpleasant (e.g., it aroused cognitive dissonance). The latter explanations would be consistent with work showing that people prefer information that confirms their views of the world (e.g., Jost, Banaji & Nosek, 2004; Nickerson, 1998). In future research, it will be important to investigate what causes participant dropout, how it can be minimised, and to carefully consider how to set timers for participants.

More broadly, it is important to realise that people can cognitively process CSTs in at least three ways (Plaks et al., 2001), as discussed in Chapter 4: People may (1) decrease their engagement with stereotype-inconsistent information and instead focus stereotype-consistent information, (2) actively engage with the counter-stereotypical information but try to debunk or re-interpret it in a way that preserves

the stereotype, or (3) respond to CSTs with an open mind by actively trying to resolve the inconsistency, as specified in the CPAG model. It was expected that the third strategy would be most likely used by participants in the present research because the CSTs were not personally relevant or threatening and medium in intensity (Gocłowska, Damian, & Mor, 2017b). However, because no measures were taken of the described processes, it is not possible to draw clear conclusions about how exactly participants reacted to the information about the target individuals. Future research is strongly advised to measure these processes (e.g., by content-coding participants' responses).

As discussed in section 4.3.3, there have recently been claims of a field-wide tendency to favour the publication of significant and/or novel effects over replications or null effects. As a result, existing findings may not represent the true population effect sizes, and this could apply to the effects of longitudinal exposure to CSTs on cognitive flexibility and intergroup bias as well. If the true population effect sizes of the predicted effects are much smaller than expected, then the power analysis conducted for the longitudinal experiment may have been misguided, and the experiment may not have been able to detect effects that are small or medium-sized. Future research is required to evaluate the reliability, validity, and generalisability of the effects of CST interventions on cognitive flexibility, intergroup bias, and concern for discrimination.

In the research reviewed at the outset of this chapter, Lai et al. (2016) concluded that, "the intervention effects were fleeting, lasting less than a couple days" and were "a testament to how the mind's prejudices remain steadfast in the face of efforts to

change them" (p. 1014). Lai and colleagues suggested that the effects may have been fleeting because: (1) effective mechanisms that can change intergroup bias have not been tested yet and thus the interventions did not target effective psychological mechanisms; (2) interventions need to be longer and more intensive, and / or (3) the interventions have been applied to the wrong populations. The authors deemed the first explanation unlikely because the chosen interventions reflected state-of-the-art knowledge on how implicit biases might be changed, and were considered likely to be successful by experts. Based on the findings reported in this chapter, one may argue that Lai et al. (2016) primarily tested interventions that aimed to modify cognitive processes, but only few of their interventions targeted affective and/or motivational processes. Could interventions that directly target motivational processes, such as interest and curiosity, be effective at reducing intergroup bias? Future theorising and research is required to answer this question. One starting point could be to develop and psychometrically validate trait and state measures of interpersonal and intergroup curiosity, which would allow to capture the hypothesised effects.

Lai and colleagues also doubted that making interventions longer and more intensive would be sufficient to elicit long-term change in intergroup bias. This is because, like the intervention in this chapter, the intensive and multi-faceted intervention developed and tested by Forscher et al. (2017) also failed to produce sustained change in both implicit and explicit bias. This suggests that simply making laboratory interventions longer or more intensive may not suffice to achieve psychological improvements. Rather, based on a review of the literature, Lai et al. (2016) proposed that prolonged everyday experiences (e.g., taking a semester-long

class on stereotypes and intergroup conflict, or participating in a cultural education programme) are needed to achieve such effects. Given the results presented in this chapter and the preceding two chapters, it does indeed appear that simple exposure to CSTs——whether occurring only once or repeatedly——is not the most effective way of improving cognitive flexibility and intergroup relations. Prolonged and intensive diversity interventions in the field, ones that are grounded in real-life settings and span months or even years, may be required to elicit sustained psychological change when adapting to social and cultural diversity.

6.6.3 Conclusion

The aim of the empirical work described in this chapter was to develop a longitudinal intervention that draws on the power of CSTs in an effort to boost cognitive flexibility and reduce explicit bias. The results of the experiment, however, raise doubt about how powerful simple exposure to CSTs is. Of course, one experiment cannot answer this question definitively (even one that was designed rigorously); thus, additional work is needed. Additional work may include theory development, the construction and validation of new measures, direct and conceptual replications, or content-coding of written responses provided by participants, to name some of the suggestions discussed earlier. The next and penultimate chapter will delineate a new theoretical model called Expectancy Regulation Theory (ERT), which aims to help specify the psychological consequences of stereotype violations, and expectancy violations more broadly, thus contributing to theory development.

Chapter 7

Expectancy Regulation Theory

As outlined in Chapter 2, it can be argued that counter-stereotypical diversity represents one type of a broader spectrum of experiences that are unexpected and surprising in nature. In an attempt to synthesise work on psychological responses to unexpected information and events, my collaborators and I have developed Expectancy Regulation Theory (ERT). ERT integrates a range of findings and existing models to specify the emotional, motivational, cognitive, and/or behavioural consequences of expectancy-violating events and derives novel and testable hypotheses. Drawing on dual process theories in other domains, ERT predicts that the basic need to understand and predict events in the environment (termed 'core prediction motive') drives a cognitive monitoring system comprising a quick, heuristic processing system (System 1, the default) and a slow, reflective processing system (System 2). ERT predicts that, whenever predictions are disrupted (e.g., when information and events violate expectancies), system 2 is triggered and, depending on

the motivation and capacity that is available to individuals, they can respond to the unexpected information in different ways. ERT predicts that (1) when people have sufficient capacity and motivation they will view the violation as a challenge to be resolved (followed by open-mindedness); (2) when they are motivated but do not have sufficient capacity, then people will view the violation as a threat (followed by defensiveness); and (3) when people lack motivation, then regardless of their capacity they will be indifferent to the violation. This chapter highlights the need to study the effects of expectancy violations over time and suggests ways to help people shift from viewing expectancy violations as threatening, to viewing them as challenging, with all of the benefits that such a reappraisal affords.

7.1 Introduction

Humans have accomplished a range of astonishing feats. They have managed to cultivate and breed animals and plants in order to produce food, clothing, and fuel. They have invented cars, computers, airplanes, and even rocket ships that fly into space. Humans' scientific curiosity, ambition, and large-scale cooperation have allowed us to eliminate infectious diseases such as smallpox and malaria and to develop antibiotics and vaccinations to keep people healthy. From Banksy to Mozart, humans create the most intellectually stimulating and emotionally uplifting art and music. And, despite the many conflicts and wars having raged in human history, 48 countries have been able to agree on a Universal Declaration of Human Rights following the Second World War, a milestone document establishing basic individual

rights such as the right to life, prohibition of slavery, and the right to freedom of thought. What has made all these achievements possible? What has motivated and enabled humans to examine life around them, to expand into new territories, and to creatively explore new ideas? One perspective on these developments is that they share an element of unexpectedness. That is, either an obstacle or challenge came to the fore that needed to be overcome, or something deeply disturbing and shocking happened that no one thought possible. For example, the Universal Declaration of Human Rights was developed in response to the unexpected atrocities committed during the Second World War, and serves to protect basic human rights irrespective of race, sex, language, religion, political opinion, or national origin.

7.1.1 Basic Psychological Needs

I propose that one key reason for the adaptive success of humans is their basic psychological need to predict events in the world, which drives their motivation and ability to interpret unexpected events as challenges rather than threats. Aristotle famously wrote in Metaphysics (350BC) that "all men by nature desire to know" and that "the human race lives by art and reasoning" (p. 1). He argued that humans proactively pursue knowledge in contrast to other animals that reactively "live by appearances and memories" (p. 1). In other words, Aristotle postulated that a basic "need to know" fundamentally drives human scrutiny and exploration (see Berlyne, 1954 and Litman & Spielberger, 2003 for more recent accounts of this argument). Fast forward to contemporary psychological science and you will

observe that researchers have converged on the idea that people have not just one, but multiple basic psychological needs (Sheldon, Elliot, Kim & Kasser, 2001). For example, Ryan and Deci (2000), like Maslow (1943) before them, proposed that there are universal human needs that are essential for health, well-being, and growth. While the exact nature and number of basic psychological needs remains contested, researchers broadly seem to agree that people have a need for: a) *acceptance / relatedness* (Ainsworth, 1979; Baumeister & Leary, 1995; Bowlby, 1969; Deci & Ryan, 2000; Dweck, 2017; Stevens & Fiske, 1995), b) *predictability* (Berlyne, 1954; Crisp, 2015; Litman, 2008; Mayseless & Kruglanski, 1987; Perlovsky & Ilin, 2012; Stevens & Fiske, 1995), c) *competence* (Dweck, 2017; Piaget, 1953), and d) *control / autonomy* (Bandura, 1982; Brehm, 1966; Dweck, 2017; Ryan & Deci, 2000). In what follows, I delve more deeply into the need for predictability, proposing that it plays a special role in human cognition and behaviour because it determines the extent to which humans can survive and thrive.

7.1.2 The Need for Predictability, the Core Prediction Motive, and Auxiliary Motives

> "Our brains have evolved to ensure our protection and survival, and to survive we need to predict." (Crisp, 2015, p. 17)

As described above, humans need to be able to understand and predict events in order to ensure their survival and well-being (Crisp, 2015; Perlovsky & Ilin, 2012; Stevens & Fiske, 1995). Here I propose that the external world——i.e., any external

stimuli, information, events, and experiences—imposes a need for predictability that, in turn, creates a need to understand and predict cause and effect, which I call the core prediction motive. In other words, the need for predictability drives the core prediction motive by propelling people to extract information from events in order to grasp their structure, causes, and consequences. Both humans and animals predict events in the world by constructing mental models involving if-then rules or contingencies (Pavlov, 1927; Skinner, 1938; Woodworth & Thorndike, 1901). This basic form of learning is essential to the survival of any living being because it allows them to adaptively respond to different cues and to effectively navigate the environment. I propose that the core prediction motive is the key motivational force driving people's ability to understand and predict events, and that it is therefore of prime importance for survival, well-being, and growth. It enables people to make sense of past events and to build new knowledge and skills to master future events. For example, experience might teach someone that if they drink too much red wine, then they will get a hangover or, if they touch a hotplate, then this will lead to burns.

Further, I suggest that the core prediction motive consists of two sub-motives: The motivation to seek meaning and the motivation to obtain knowledge. First, people may ask why and how an event happened, and thus construct more elaborate models about how the world works (e.g., prototypes, stereotypes, ideologies, culture). I refer to this as the motivation to seek meaning. Second, humans have the, relatively unique, capacity to develop complex models of how the world works (e.g., more complex if-then rules, mathematical models of the world), which I refer

to as the motivation to obtain knowledge. The meaning and knowledge motives significantly differ from each other in nature: While the motivation to seek meaning tends to be biased, the motivation to seek knowledge tends to be accuracy-driven, as I elaborate next.

7.1.3 Biased versus Accuracy-driven Information Processing

I hypothesise that the core prediction motive can at times be biased and at other times accuracy-driven, depending on which motive it is driven by (i.e., to seek meaning or obtain knowledge). Although a significant amount of psychological research suggests that people tend to engage in motivated reasoning that serves their expectancies and helps to protect their ideologies (Jost, Glaser, Kruglanski & Sulloway, 2003; Kunda, 1990; Nickerson, 1998), it is important to recognise that individuals are also often motivated to engage in accurate and data-driven information processing (Clark, 2013; Stevens & Fiske, 1995). Here I propose that the motivation to seek meaning is a drive for an internal state or feeling that reality is ordered and unambiguous. In contrast, the motivation to obtain knowledge refers to the external world and the desire to make accurate predictions of future events. For example, if an aspiring entrepreneur desperately clings to a startup business idea despite abundant evidence suggesting that it is not feasible (e.g., potential investors declining to invest or the product failing to grow in the relevant market), then at some point reality will force the entrepreneur to face his or her situation

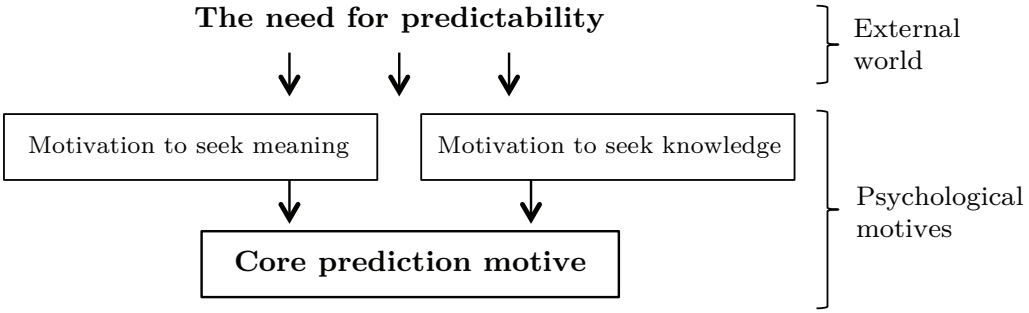

Figure 7.1: Overview of the proposed motives constituting the core prediction motive: the meaning motive and the knowledge motive.

objectively. He or she will have to start making decisions that are based on reality (i.e., he/ she will need to make sure that he can make a living) rather than based on his or her subjective (and thus biased) hopes or dreams (e.g., his/her desire to launch a trailblazing startup). Thus, I suggest that while the biased, defensive style of information processing (driven by the meaning motive) may produce a sense of protection and reduce feelings of uncertainty in the short-term, it is open-minded, unprejudiced cognitive processing (driven by the knowledge motive) that allows people to better understand and predict events in the long-term. In sum, I propose that, under the right circumstances, the core prediction motive can powerfully catalyse accuracy-driven (rather than biased) cognitive processes. An overview of the proposed motives constituting the core prediction motive is presented in Figure 7.1.

7.1.4 Psychological Reactions to Unexpected Information and Events

As people pursue goals that fulfill their needs (e.g., a student reads a book to acquire knowledge), they build mental representations of their experiences that represent their expectations or beliefs about the world (e.g., said student learns that women are underrepresented in the tech industry; Dweck, 2017). However, these expectations and beliefs are not always confirmed. For instance, as discussed in the preceding chapters, in today's increasingly socially and culturally diverse societies new, cross-cutting forms of diversity are emerging that disconfirm and even erode traditional stereotypes. There are examples of female chief executives spearheading big technology firms, Muslims advocating liberal values such as abortion rights, disabled people winning sports competitions, Catholics coming out as gay, or successful startup founders who dropped out of university. Other examples of events that disconfirm expectations include politicians or political parties surprisingly winning elections despite polling forecasts predicting otherwise (e.g., the election of Donald Trump as president of the USA in 2016), computer programs surprisingly beating humans at board games (e.g., in 2016 AlphaGo became the first computer program to beat a human professional Go player), financial or economic crises unexpectedly breaking loose (e.g., the financial crisis of 2007-2008), or unexpected natural catastrophes wreaking havoc (e.g., the 2017 Atlantic hurricane season shattering many Caribbean islands).

For many decades, scientists have documented the aversive psychological and

physiological consequences of events or stimuli that violate expectations (Batson, 1975; Bruner & Postman, 1949; Cooper & Jahoda, 1947; Elliot & Devine, 1994; Harmon-Jones, 2000; Killian, Festinger, Riecken & Schachter, 1957; Plaks, Grant & Dweck, 2005; Proulx, Inzlicht & Harmon-Jones, 2012; Topolinski & Strack, 2015). Across these paradigms, researchers have viewed expectancy violations as precursors to stress, negative affect, and defensive cognitive processing. Some also have argued that expectancy violations evoke threat-relieving responses, which involve an array of palliative compensation efforts, such as assimilation (reinterpreting experiences to be consistent with expected relationships) or affirmation (committing to alternative expected relationships; Proulx et al., 2012).

I propose that this focus on the potentially threatening nature of unexpected events and stimuli illuminates only one possible—albeit important—mechanism through which people respond to expectancy violations. Specifically, I aim to supplement the traditional and somewhat negative research focus with a more recent, more positive perspective, which is that, under certain circumstances, expectancy violations can be powerful drivers of positive affect and enhanced cognitive flexibility (Crisp & Meleady, 2012; Crisp & Turner, 2011; Gocłowska et al., 2017a; Gocłowska et al., 2017b; Ritter et al., 2012; Ritter et al., 2014). Understanding when and why people respond positively versus negatively to expectancy violations can help to provide some answers to important questions such as: How do humans respond and adapt to deviations from the status quo? And how can we promote more adaptive behaviours in response to unexpected information and events?

In an effort to provide this understanding and answer these questions, I propose

an integrative theory based on a synthesis of a variety of research programs ranging from work in psychophysiology (e.g., Blascovich, Mendes, Tomaka, Salomon, & Seery, 2003; Tomaka, Blascovich, Kibler, & Ernst, 1997), affective science (e.g., Bartholow, Fabiani, Gratton, & Bettencourt, 2001; Bettencourt & Manning, 2016; Gross, 1998); motivation science (e.g., Kleiman & Hassin, 2013; Litman, 2008; Litman & Spielberger, 2003), positive psychology (e.g., Fredrickson, 2001; Fredrickson & Branigan, 2005), cognitive science (e.g., Noordewier, Topolinski, & Dijk, 2016; Proulx et al., 2012; Ritter et al., 2012), diversity science (e.g., Crisp & Turner, 2011, Crisp & Meleady, 2012, Gocłowska, Damian, & Mor, 2017b, and personality psychology (e.g., Gocłowska, Baas, Elliot, & De Dreu, 2017a). My integrative theory suggests that expectancy violations have much broader affective, motivational, cognitive, and behavioural effects than previously assumed and can, in many cases, have positive outcomes. The integrative theory, which I term Expectancy Regulation Theory (ERT), is summarised in Figure 7.2.

The remainder of this review is presented in four parts. First, expectancies and expectancy violations are defined and disambiguated, in order to locate these constructs in the context of related ideas and psychological theories. Second, the key hypotheses derived from ERT are described and current evidence in support of the theory, as well empirical gaps to be addressed, are summarised. Third, the ERT's theoretical relationship with existing models and theories is critically discussed. Lastly, directions for future research on the psychological adaptation to unexpected information and events are proposed.

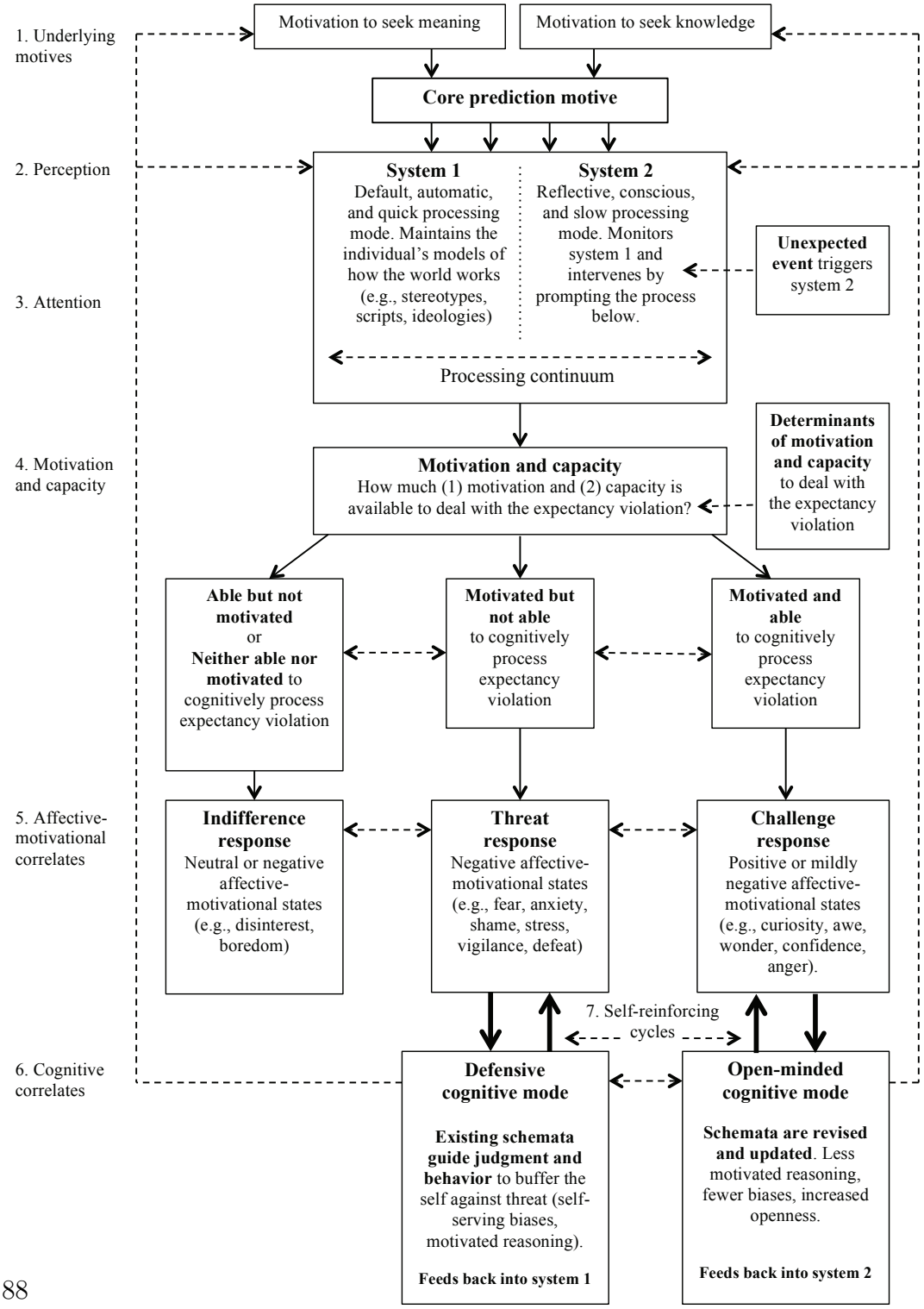

Figure 7.2: Schematic presentation of Expectancy Regulation Theory (ERT). Dotted arrows represent flexible pathways that can be influenced by individual or contextual variables.

7.2 Definitions and Disambiguations

7.2.1 Definitions

Expectancies can be defined as assumptions or hypotheses about the world that guide effective behaviour. For example, stereotypes are expectancies about social groups, and while they may or may not accurately reflect reality (Judd & Park, 1993), they are heuristics that help individuals cognitively organise their experiences (Fiske & Taylor, 1991; Kahneman & Tversky, 1973). Expectancies may be general (e.g., pertaining to societal norms or general physical laws such as gravity) or particular (e.g., pertaining to specific individuals or interactions; Burgoon, 1993).

Expectancy violations are stimuli or events that contradict people's assumptions about the world. Specifically, I define expectancy violations as events or stimuli that deviate from enduring (i.e., global) or activated (i.e., situational) cognitive schemata (Schützwohl & Reisenzein, 1999). Put differently, expectancy violations are information or events that people did not predict and they can thus create uncertainty (Mendes et al., 2007). Expectancy violations are related to constructs like cognitive dissonance (Festinger, 1957), uncertainty (Hogg, 2007; McGregor, Zanna, Holmes & Spencer, 2001), ambivalence (Nordgren, van Harreveld & van der Pligt, 2006; Thompson, Zanna & Griffin, 1995), and novelty (Berlyne, 1954; Förster, Marguc & Gillebaart, 2010), yet they are conceptually distinct. In what follows, I briefly review and disambiguate these related constructs.

7.2.2 Disambiguations

7.2.2.1 Cognitive Dissonance

In 1957, Leon Festinger published A Theory of Cognitive Dissonance and sparked a wealth of research. The key tenet of the theory holds that inconsistent but mutually relevant elements of knowledge (i.e., "cognitions") produce a state of discomfort (termed "dissonance"), which people strive to alleviate. Festinger conceptualised "cognitions" broadly, as "any knowledge, opinion, or belief about the environment, about oneself, or about one's behaviour" (Festinger, 1957, p. 3). The notion of inconsistent cognitions is conceptually equivalent to the construct of expectancy violations because both involve a contradiction between at least two cognitions about oneself or the world; however, the idea of expectancy violations is broader than that of inconsistent cognitions because it can refer to unexpected events and stimuli that occur outside the person——e.g., meeting a Harvard-educated carpenter. However, dissonance is clearly one consequence of expectancy violations, and there is evidence to suggest that it can be reduced by engaging in defensive (Proulx et al., 2012; Randles, Inzlicht, Proulx, Tullett & Heine, 2015) as well as open-minded processes (e.g., McFalls & Cobb-Roberts, 2001).

7.2.2.2 Uncertainty

According to Wilson (1973), uncertainty is a feature of all situations that involve innovation, novelty, ambiguity, complexity, or risk. Shafir and Tversky (1992) provide a narrower definition of uncertainty as "disjunctions of possible states: either

one state will obtain, or another" (pp. 449–450). In yet another definition, Downey and Slocum (1975) postulate that uncertainty can be defined as "a state that exists when an individual [engages] in directed behaviour based upon less than complete knowledge (...)" (p. 571). These different definitions seem to contain one common element: Uncertainty implies that there is a discrepancy or gap in the available information. As such, uncertainty can, but need not follow from expectancy violations. For example, if I grew up being poor, then I might experience a persistent sense of uncertainty, but this type of uncertainty will be relatively predictable and therefore not constitute a violation of my expectancies.

7.2.2.3 Ambivalence

Ambivalence is defined as the co-existence of positive and negative evaluations of an attitude object (Nordgren et al., 2006). Because ambivalence constitutes a response to an attitude object, and because attitude objects can violate expectancies, it follows that people can experience ambivalence with respect to attitude objects that violate their expectancies. For example, if someone lives in a democratic country and thinks that democratic processes should be promoted (e.g., voters should be educated about politics), but in the most recent election some democratic principles were violated (e.g., it was found that some politicians had manipulated the election process), then this person may feel ambivalent about the state of democracy in their country. Ambivalence is conceptually similar to cognitive dissonance and uncertainty because it is preceded by a cognitive discrepancy——the difference is that, while ambivalence follows from a discrepancy within an attitude (i.e., an intra-

attitudinal discrepancy), cognitive dissonance and uncertainty often result from discrepancies between attitudes (van Harreveld, Nohlen & Schneider, 2015) or between attitudes and behaviour (e.g., a person might have positive attitudes toward democracy, but not vote), or the outcomes of behaviour (the candidate who wins an election was not expected to do so). Also, note that people can be ambivalent about a topic and at the same time be certain about this ambivalence, which supports the idea that ambivalence and uncertainty are distinct constructs.

7.2.2.4 Novelty

Lastly, the concept of novelty is conceptually related to the idea of expectancy violations; however, it is not equivalent to it. Novelty is commonly defined as something that has not been previously experienced and therefore something that people are not familiar with (e.g., see Förster, Marguc, & Gillebaart, 2010; Zajonc, 1998). But experiences do not have to be novel to violate expectancies. For example, someone might be aware that men can work as midwives or kindergarten teachers, yet they may still be surprised to actually meet a male midwife at a hospital. This means that expectancy violations are context-specific, such that expectations might get challenged even if people have knowledge about a domain (i.e., familiarity).

Taken together, the preceding sections helped establish conceptual clarity and theoretical boundaries by (1) defining expectancy violations and (2) specifying how they are similar to and distinct from related constructs.

7.3 Hypotheses Derived from ERT

Hypothesis 1: The core prediction motive supersedes other motives.

The first key hypothesis of ERT is that the core prediction motive is an ever-present motivational force so critical to human cognition and behaviour that it supersedes other motives, but at the same time is fuelled by secondary motives (namely, the motivations to seek meaning and obtain knowledge; see "1. Underlying motives" in Figure 7.2). Because the very survival of any organism depends on the ability to predict events (Clark, 2013; Hohwy, 2014), I suggest that the core prediction motive is even more fundamental than other needs, such as the needs for acceptance, connectedness, and competence. More precisely, the needs for acceptance and/or competence are highly unlikely to manifest themselves without a basic need to predict. For example, if a person cannot reliably link basic cause and effect, then he or she cannot know whether certain behaviours (e.g., hard work) are likely to promote acceptance or competence. As a result, striving for acceptance or competence without the ability to predict likely outcomes will quickly become futile. Therefore, the core prediction motive is predicted to supersede other basic psychological motives and needs. Future research could test this hypothesis by using measures like the EMMM (exploratory measure of multiple motives; Rubin, 2017), which allows to juxtapose a range of psychological motives and to draw inferences about the relative importance of different motives for psychological phenomena.

Hypothesis 2: The core prediction motive drives the cognitive monitoring

system.

According to ERT, the core prediction motive fuels a cognitive monitoring system comprising two sub-systems (see "2. Perception" in Figure 7.2). Drawing on dual process theories in other domains (Crisp & Meleady, 2012; Evans, 2008; Kahneman, 2012; Kahneman & Tversky, 1973; Strack & Deutsch, 2004), I propose that these two sub-systems, termed system 1 and system 2, operate on a processing continuum ranging from quick and heuristic processing to slow and reflective processing. More precisely, system 1 represents the default mode of processing that helps people to maintain their models of how the world works (e.g., stereotypes, scripts) by operating in an automatic and unconscious fashion. In contrast, system 2 represents a slower, more deliberative and conscious mode of processing that helps people to understand stimuli and events, especially those that do not conform to existing models. Typically, system 1 operates when people engage in routine activities, such as reading simple sentences or driving a car on an empty street. System 2 operates when more focused attention is needed, for example when studying for an exam or trying to understand why a friend is upset.

Some evidence for the idea that the core prediction motive is a driving force of cognitive processing can be found in early research testing attribution theory (Heider, 1958; Heider & Simmel, 1944; Jones & Davis, 1965; Kassin, 1979; Kelley, 1967; Leary et al., 2017; Mcarthur, 1972). For example, Heider and Simmel (1944) asked participants to interpret a film showing geometric figures moving in various directions and at various speeds. The researchers found that participants generally

described the movements in terms of human intentions and motives, thus making interpretations of cause and effect rather than merely describing the events. While reviewing research on attribution theory is beyond the scope of this chapter, it shall suffice to say that extant research demonstrates that people have an intrinsic need to understand and explain the causes of behaviour and events (e.g., Kelley, 1973; Kelley & Michela, 1980). Future research needs to test whether it is the core prediction motive that drives the cognitive monitoring system (i.e., the operation of systems 1 and 2), or whether other basic motives and/or needs are involved.

Hypothesis 3: System 2 monitors system 1.

The third hypothesis derived from ERT is that system 2 monitors system 1 and overrides it when necessary, for example when an unexpected event happens or new and surprising information cannot be readily reconciled with existing knowledge (see "3. Attention" in Figure 7.2). In other words, system 2 helps people to realise that one or more of their expectations have been violated, or that they have made a prediction error. For example, if someone travels to a new culture (e.g., India) and tries to indicate their polite refusal to an offer by shaking their head (as is common in Western cultures), they are likely to discover that their head shaking has the opposite effect to the one that they intended. This is because in India headshaking is considered an affirmation, so it will be assumed that the traveller is accepting the offer even though they meant to refuse it. Only by experiencing the unexpected consequences of their actions will the traveller be able to realise that their initial expectancies have been violated, leading system 2 to be activated. If, however, there

is no explicit or implicit feedback then they are unlikely to realise that their original assumption was misguided, and so they will likely continue to operate based on her original assumption(s)—i.e., in accordance with system 1.

The hypothesis that the perception of, and attention to, the expectancy violation is necessary in order for it to shift the mode of processing is consistent with a host of research showing that expectancy violations invoke physiological arousal that alerts the individual to the unexpected stimulus or event (Burgoon, 1993; Dickter & Gyurovski, 2012; Jussim, Coleman, & Lerch, 1987; Norman & Shallice, 1986; see also Sokolov, Spinks, Näätänen, & Lyytinen, 2002 for an overview of research on the orienting response). Furthermore, the prediction that an overarching system monitors the relation between the world and the person's expectancies and signals discrepancies (i.e., events and stimuli that violate these expectancies) is consistent with research on the experience of surprise, a marker of the extent to which expectancy violations command attention (Horstmann, 2015; Noordewier & Breugelmans, 2013; Schützwohl & Reisenzein, 1999). Finally, this hypothesis is in line with developmental research on the experience of surprise. For example, research by Camras et al. (2002) showed that infants were judged to be more surprised and interested when they were presented with an event that violated their expectations (i.e., when a toy was covertly switched) as compared to an event that they had previously experienced and presumably came to expect. In sum, ERT predicts that when unexpected information is encountered, system 2 initiates a process that, at its best, helps challenge and revise existing intuitions (see also Meleady & Crisp, 2014). But how exactly does it do so?

Hypothesis 4: Motivation and capacity determine psychological responses to expectancy violations.

The fourth hypothesis derived from ERT is that, upon having noticed the expectancy violation, individuals' (1) motivation and (2) capacity determine how they deal with the expectancy violation (see "4. Motivation and capacity" in Figure 7.2). The extent to which an expectancy violation catalyses somebody's motivation to deal with the event can depend on many different individual and contextual variables, and interactions thereof. In the following, I decompose this general hypothesis into three more specific sub-hypotheses.

Hypothesis 4a: People need to be (appropriately) motivated in order to effectively deal with expectancy violations.

Given that the core prediction motive is assumed to trump all other possible motives, I reason that people's motivation to deal with an expectancy violation is primarily determined by the extent to which it disrupts their predictive ability. That is, if a new piece of information or an unexpected event implies that a person's working model of the world is substantially flawed and therefore presents a threat to their ability to navigate their environment, then it is likely that they will be motivated to cognitively process this expectancy violation. Thus, I hypothesise that people may be particularly motivated to attend to and process expectancy violations indicating that they may lose something important to them. In the case of CSTs, one would not typically expect that they significantly disrupt people's predictive ability

because most of the time they represent conjunctions of *already known categories*. It appears that entirely novel and unexpected categories, developments, or events (e.g., space travel) are most likely to motivate people to pay attention to expectancy violations. However, the results from Chapter 5 on NFC as a moderator of the effects of exposure to CSTs on cognitive reflection raise a caveat: They seem to indicate that too much cognitive motivation may potentially be counter-productive. Thus, because there seems to be a desirable 'sweet spot' that may be most conducive to coping with expectancy violations, it is suggested that people need to be *appropriately* motivated in order to effectively deal with expectancy violations.

Hypothesis 4a is broadly supported by decades of research on primary appraisal in potentially stressful situations, suggesting that individuals first evaluate the nature and degree of risk that a situation presents (Lazarus & Folkman, 1987; Tomaka, Blascovich, Kelsey & Leitten, 1993). A wealth of research has further demonstrated that threatening stimuli and outcomes are surprising and thus help prioritize attention (Mogg et al., 2000; Notebaert, Crombez, Van Damme, De Houwer, & Theeuwes, 2011; Schützwohl & Borgstedt, 2005; for an overview see Jonas et al., 2014). That is, attending to events that violate expectancies and are potentially threatening is of obvious importance to survival and well-being.

Considerable research also suggests that losses are psychologically more powerful than gains (Kahneman, Knetsch, & Thaler, 1991; Tversky & Kahneman, 1991; see also Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001), lending support to Hypothesis 4a. To give a concrete example, typical experiments examining loss aversion show that people generally prefer avoiding losses (e.g., losing $5) as compared to

making equivalent gains (e.g., gaining $5). Taken together, it follows that the more an expectancy violation indicates that people may lose something they value, the more motivated people will be to deal with it.

> *Hypothesis 4b: People need to have sufficient capacity in order to effectively deal with expectancy violations.*

In addition to people's motivation to process the expectancy violation, I predict that people's resources, or capacity, to deal with the expectancy violation determines their subsequent psychological responses. Here, capacity refers to sufficient resources for executive functioning (i.e., higher order cognitive operations that are involved in the planning, execution, and regulation of behaviour; Baddeley, 1996; Baddeley & Hitch, 1974; Norman & Shallice, 1986). While the core prediction motive is the key driver behind people's motivation to make sense of unexpected information or events, it less obvious how it relates to their capacity or the resources that they have available to process the expectancy violation. I propose that the core prediction motive is largely independent of people's capacity to cope with unforeseen events, yet this does not detract from the general importance of the person having capacity or adaptive resources at their disposal when dealing with unexpected information or events.

Research on appraisal theory (Lazarus, 1991; Lazarus & Folkman, 1987) supports this idea, by stressing the importance of adaptive resources (e.g., good health, problem-solving skills, positive beliefs, social skills, social support, material resources) in the face of stressors. For example, imagine that someone unexpectedly

loses their job, which substantially undermines their ability to plan and predict their life. Arguably, they would have at least the following three options: they may (1) feel indifferent (e.g., the person did not enjoy the job or suffers from depression), (2) feel threatened and become defensive (e.g., trying to save their job somehow), or (3) if they have sufficient capacity they may feel challenged and try to find creative ways out of the precarious situation (e.g., by applying for new jobs). If people lack the physical resources to engage in executive functioning, then they will not be able to plan and regulate their behaviour effectively in response to unexpected information and events. Thus, people's capacity to deal with expectancy violations is hypothesised to be a critical determinant of how they cope with unexpected information or events. Future research is needed to test the reliability and generalisability of this predicted effect in diverse contexts and samples.

> *Hypothesis 5: Indifference versus threat versus challenge may follow depending on the capacity and motivation available to deal with the expectancy violation.*

The fifth hypothesis derived from ERT is that expectancy violations can prompt one of three responses—indifference, threat, or challenge (see "5. Affective-motivational states" in Figure 7.2). Whereas the threat response is characterised by negative affective-motivational states (e.g., feelings of fear, stress, shame, defeat), the challenge response involves positive or mildly negative affective-motivational states (e.g., feelings of curiosity, awe, wonder, confidence), and an indifference response tends to involve neutral or negative affective-motivational states (e.g., feel-

ings of disinterest, boredom). ERT proposes that, which of the three responses is triggered depends on individuals' motivation and capacity to deal with the expectancy violation, which in turn are determined by certain individual difference variables and/or situational circumstances. Below (under Hypothesis 6), I review examples of moderators in order to help explain when people respond to unexpected events with indifference, a threat response, or a challenge response. However, for now the processes described can be distilled into three predictions (5a, 5b, and 5c).

> *Hypothesis 5a: An indifference response and neutral/negative emotions can follow as a result of expectancy violations.*

ERT predicts that insufficient motivation to deal with an expectancy violation (regardless of the person's capacity to do so) will typically result in indifference. Specifically, individuals who are able but not motivated, or neither motivated nor able, to make sense of the expectancy-violating event, will show an indifference response and experience neutral or negative affective-motivational states (e.g., disinterest, distraction, boredom). While I am not aware of research that directly tests this prediction, emerging research on some of the hypothesised indifference states, such as boredom, implicitly supports ERT. For example, research on the experience of boredom suggests that "feeling unchallenged" and perceiving activities as "meaningless" are central characteristics of boredom, setting it apart from other emotions and implying that the bored person is typically not sufficiently engaged with his or her present situation (van Tilburg & Igou, 2012). Future research needs

to test directly whether expectancy violations under certain conditions can trigger an indifference response.

> *Hypothesis 5b: A threat response, negative emotions, and defensive cognitive processing can follow as a result of expectancy violations.*

ERT further predicts high motivation but low capacity to deal with an expectancy violation will likely result in a threat response. More precisely, individuals who are motivated but not able to make sense of the expectancy-violating event will experience negative affective-motivational states (e.g., fear, stress, vigilance, shame) and ultimately, engage in cognitively defensive cognitive processing (e.g., assimilation or affirmation of existing meaning frameworks; Heine, Proulx, & Vohs, 2006; see "6. Cognitive correlates" in Figure 7.2). More precisely, I reason that, in such a context, existing schemata will predominantly guide individuals' judgment and behaviour (Crisp & Turner, 2011), which serves the goal of buffering the self against the threat by any means possible.

Indeed, a substantial body of research, synthesised by the meaning maintenance model (Heine, Proulx & Vohs, 2006; Proulx & Inzlicht, 2012), suggests that people naturally seek to generate and maintain a sense of meaning (defined as a system of expected relations, Heine, Proulx, & Vohs, 2006). However, if their sense of meaning is threatened, then they will seek to reaffirm meaning in related or even unrelated domains. For example, it has been found that people who are reminded of their mortality are more prejudiced against unrelated outgroups (McGregor et al., 2001), protect cultural icons (Greenberg, Porteus, Simon, Pyszczynski & Solomon, 1995),

and start believing more intensively in supernatural agents (Norenzayan & Hansen, 2006). Furthermore, extensive research on self-affirmation suggests that threats to the self, such as negative social comparisons or cognitive dissonance, can be offset by self-affirming one's values (Steele & Liu, 1983; Tesser, 2000; Tesser, 2001; Tesser & Cornell, 1991).

While there has been extensive research on physiological threat responses and the affective states that they trigger (Blascovich, Mendes, Tomaka, Salomon & Seery, 2003; Tomaka et al., 1993; Tomaka, Blascovich, Kibler & Ernst, 1997), there has been very little research directly examining the question whether or not unexpected events, by triggering threat states, lead to defensive processing. One rare example of such research is the work by Mendes et al. (2007), who examined the physiological responses during social interactions with people who violated expectations. In the first two experiments, participants either interacted with White or Latino confederates who either confirmed or contradicted common stereotypes (i.e., they were described as either high or low on socio-economic status, SES). The researchers found that participants who interacted with counter-stereotypical partners (i.e., Latino partners high in SES, or Whites low in SES) exhibited cardiovascular responses consistent with threat, they showed poorer task performance, and they manifested negative and defeat-related behaviour. In a third experiment, Mendes and colleagues conceptually replicated this finding by having participants interact with Asian or White confederates who spoke with expected or unexpected accents. Again, participants who interacted with counter-stereotypical individuals exhibited a physiological threat response and poorer task performance as compared to par-

ticipants who interacted with stereotypical partners. Remarkably, the same effect was found when the counter-stereotypical partner was positively surprising. This research lends support to the idea that a physiological threat response will likely lead to defensive cognitive processing. However, these experiments have several limitations. The conclusions drawn from the work by Mendes et al. (2007) are constrained to the extent that the experiments included tasks that constituted a motivated performance situation——i.e., tasks that were goal-relevant and required instrumental cognitive responses. This may have put some demand on participants and consumed cognitive resources from the outset and as a result, participants may potentially have become anxious about their performance. It then seems plausible that——on top of this baseline demand——participants experienced an interaction with a counter-stereotypical stranger as threatening, rather than challenging. Participants might also have felt the need to compete with their interaction partners while working on the tasks, which may have further increased the likelihood that they felt threatened. Future research is required confirm the replicability and generalisability of the finding that expectancy violations can prompt a threat response, which in turn leads to cognitively defensive processing.

> *Hypothesis 5c: A challenge response, positive emotions, and open-minded cognitive processing can follow as a result of expectancy violations.*

What happens when individuals are sufficiently motivated and able to engage with the expectancy-violating event? ERT predicts that, under these circumstances,

expectancy violations can lead to a psychological challenge response, which entails positive or mildly negative affective-motivational states (e.g., interest, curiosity, awe, wonder, confusion) and ultimately leads to cognitively expansive cognitive processing (see "6. Cognitive correlates" in Figure 7.2; Crisp & Turner, 2011; Fredrickson, 2001; Gocłowska, Baas, Elliot, & De Dreu, 2017a; Kleiman & Hassin, 2013; Ritter et al., 2012). That is, the individual will be more likely to update existing schemata and generate new ones in order to make sense of the expectancy violation (Crisp & Turner, 2011), which can even take place in unrelated domains. In support of this idea, Ritter et al. (2012) explored the effects of diversifying experiences (defined as the active involvement in an unusual event) on cognitive flexibility. In one experiment, Ritter and colleagues found that actively experiencing complex, unusual, and unexpected events in a virtual reality environment increased participants' cognitive flexibility in an unrelated creativity task. In a second experiment, Ritter and colleagues observed that participants who experienced minimal schema violations (i.e., the reversal of steps of a simple everyday activity like preparing a sandwich) again displayed enhanced cognitive flexibility on an unrelated creativity task. Similarly, in cross-cultural contexts it has been found that people who were reminded of disorienting cultural experiences while living abroad demonstrated improved creative problem solving (Maddux, Adam & Galinsky, 2010) and research has demonstrated that people who were reminded of their own mortality drew T-shirt designs that were judged to be of higher artistic quality——but only if they had a low need for structure (Routledge & Juhl, 2012). Lastly, two studies examining the cognitive effects of awe demonstrated that participants who had just completed an

awe-induction task (awe was defined as an experience that challenges the person's frame of reference, akin to expectancy violations), were less persuaded by messages containing weak arguments (Griskevicius, Shiota & Neufeld, 2010). Overall, it seems clear that expectancy-violating experiences can often promote systematic processing, cognitively expansive processing, and even the "wholesale assembly" of new meaning frameworks (Proulx & Inzlicht, 2012). But while there is work on the effects of expectancy violations on cognitive processes (e.g., cognitive flexibility and creativity), I am not aware of any work directly linking psychological challenge states to cognitive flexibility and creativity. In part this is because the literature on expectancy violations has predominantly examined whether or not expectancy violations lead to threat, and disregarded the possibility that they may lead to a challenge response. Future research should directly test the idea that when there is sufficient motivation and capacity to deal with an expectancy violation, people can experience a challenge response and consequently engage in open-minded cognitive processing.

Lastly, recall the research by Mendes et al. (2007) described earlier, which featured participants who interacted with counter-stereotypical individuals. What would happen if, for example, the competitive nature of the experimental situation was reduced (by removing the cognitive performance task), and if threat versus challenge states were measured while participants were getting to know each other? ERT would predict that this experimental setup would engender interest and curiosity in participants who interact with expectancy-violating partners (as compared to those who interact with expectancy-confirming partners), and therefore promote

a challenge response. Future research is needed to test this idea.

> *Hypothesis 6: Different traits and states can determine whether indifference versus threat versus challenge follow as a result of expectancy violations.*

ERT proposes that the determinants of indifference versus threat versus challenge in response to expectancy violations can be related to individual trait and state differences as well to contextual or situational differences. In this section, I review individual differences in epistemic motivation (i.e., need for structure/closure and need for cognition) in order to provide empirical examples of how these variables can potentially be determinants of indifference versus threat versus challenge.

Individual differences in epistemic motivation, more specifically personal need for structure / closure and need for cognition, have been shown to play an important role in people's responses to expectancy violations. Need for structure refers to individual differences in the tendency to create and use abstract mental representations (e.g., schemata, scripts, attitudes, and stereotypes) that are simplified generalisations of previous experiences (Neuberg & Newsom, 1993). Need for closure is closely related and is defined as the motivation to achieve "an answer on a given topic, any answer, (...) compared to confusion and ambiguity" (Webster & Kruglanski, 1994, p. 1049). Because need for structure and need for closure are highly correlated (e.g., Leone, Wallace, & Modglin, 1999) I will use the terms interchangeably. Need for cognition is defined as an individual's tendency to engage in and enjoy effortful cognitive activity (Cacioppo et al., 1996) and is conceptually distinct from need for

structure / need for closure. For example, Neuberg and Newsom (1993) report only a weak, negative correlation between need for cognition and need for structure, and therefore conclude that these constructs are orthogonal. More precisely, need for cognition represents preferences for the amount of cognitive activity, whereas need for structure represents preferences for the desired outcome of any given cognitive activity. Both individual differences are likely to shape how people respond to expectancy violations.

> *Hypothesis 6a: Need for structure / need for closure moderates the psychological effects of expectancy violations*

The role of need for structure / need for closure in shaping people's responses to events and stimuli that violate their expectancies has been investigated in various programmes of research. For example, Vess, Routledge, Landau, and Arndt (2009) tested the effect of mortality salience on people's perceptions of life's meaning. The researchers found that individuals high (vs. low) in need for structure viewed life as more meaningful after being reminded of death and a tendency to defend their belief systems. In contrast, participants low (vs. high) in need for structure perceived life to be less meaningful after being reminded of their own mortality. Importantly, people low in need for structure resolved the lack of meaning by being more willing to explore novelty (e.g., increased interest in documentaries that presented alternative perspectives to culturally relevant topics), which in turn led them to affirm the meaningfulness of their own lives. Through the lens of ERT, these responses closely resemble the psychological threat versus challenge responses outlined earlier:

That is, while participants high in need for structure displayed defensive reactions and thus presumably felt threatened, it appears that participants low in need for structure, given their openness to exploring novelty, displayed a challenge response.

In a different line of research, Gocłowska, Baas, Crisp, and De Dreu (2014) examined the influence of social schema violations (e.g., a female mechanic, an astronaut on the beach, an Eskimo in a desert) on divergent thinking (defined as flexible switching among a broad range of categories). The authors discovered that exposure to social schema violations promoted creativity among participants low in need for structure, but impeded creativity among participants high in need for structure. Again, this finding is consistent with ERT because it appears that individuals high in need for structure experienced either indifference or psychological threat, whereas individuals low in need for structure appeared to experience psychological challenge. However, note that Gocłowska et al. did not measure indifference/threat/challenge states in their studies, so this interpretation remains speculative.

Consistent with the above findings, Kruglanski, Pierro, Mannetti, and De Grada (2006) summarise evidence showing that a high need for closure leads to in-group favouritism, rejection of deviates, and resistance to change. In one experiment, Kruglanski and Webster (1991) tested the hypothesis that individuals high in need for closure would be more likely to reject the opinions of those who threatened to undermine group consensus. They manipulated need for closure, and asked groups of university students to reach consensus on a case involving compulsory drug testing (at the time, the majority opinion was in favour of drug testing for campus athletes). One confederate advocated the majority opinion, while another confederate

argued against testing. The results of this and follow-up experiments indicated that increased need for closure resulted in a stronger tendency to downgrade or reject the opinion that deviated from the consensus. Thus, it appears that individuals high in need for closure are more easily disturbed when social norms are violated (see also Fu et al., 2007).

In sum, the findings described above lend indirect support to ERT. They tentatively suggest that (chronic or temporarily induced) high need for structure / closure may predispose individuals to respond to expectancy-violating events and stimuli with indifference or psychological threat, which may result in negative affective-motivational states and defensive cognitive processing. In contrast, individuals low in need for structure / closure appear to view events and stimuli that violate their expectations as a psychological challenge with its concomitant positive affective-motivational states (e.g., openness) and cognitively expansive cognitive processing (e.g., more creativity). Future research needs to directly test the proposed moderated effects by measuring indifference versus threat versus challenge states.

> *Hypothesis 6b: Need for cognition moderates the psychological effects of expectancy violations.*

Recent research has examined the role of need for cognition on the cognitive flexibility of individuals who are exposed to counter-stereotypes (see Chapter 5). Here, it was predicted that exposure to counter-stereotypes (defined as unusual and unexpected combinations of social categories) should prompt individuals to think "outside the box" and make them more cognitively flexible, but only if there is

sufficient room to be surprised and intrigued (individuals with a low need for cognition represent one such group because they do not typically think about novel or surprising events). Consistent with this idea, it was found that people low in need for cognition showed increased cognitive flexibility after being exposed to CSTs (relative to a control condition in which participants were exposed to stereotypical stimuli), whereas people high in need for cognition showed decreased cognitive flexibility after being exposed to CSTs. From the perspective of ERT, this finding suggests that participants with a low need for cognition experienced a psychological challenge response, allowing them to engage in open-minded cognitive processing, while participants with a high need for cognition experienced self-regulatory fatigue / cognitive depletion (Inzlicht & Schmeichel, 2013), which may resemble an indifference or a threat response. Again, however, the authors did not measure indifference versus threat versus challenge states directly, and so this represents an important direction for future research.

Two further insights from this work are worth highlighting in the present context. First, in Chapter 5 the importance of the intensity of the counter-stereotypical experiences in determining whether beneficial versus adverse psychological outcomes result from the experience were explicitly discussed. Specifically, it was proposed that expectancy violations that are low-to-moderate in intensity (such as imagining a counter-stereotypical individual) are likely to be experienced as challenging rather than threatening because they require fewer resources for adaptation. In contrast, expectancy violations moderate or high in intensity (such as a highly charged intergroup encounter) are likely to be experienced as threatening because they re-

quire more resources in order to be cognitively reconciled (for a similar prediction, see Gocłowska et al., 2017b). Second, Chapter 5 discovered that participants high in need for cognition displayed lower cognitive performance after being exposed to CSTs. However, a positive mediation effect was also found, such that participants high in need for cognition experienced a psychological state characterised by surprise, perceived expectancy violation, and cognitive complexity, which in turn helped boost their performance (similar to participants low in need for cognition). This result led the authors to speculate that two competing psychological processes may be at work for individuals high in need for cognition: Specifically, they may feel intrigued and cognitively challenged, but also (in part, as a result) become cognitively depleted over time. However, the mechanism underlying the positive indirect effect appeared insufficient to compensate for the negative direct effect on performance. Taken together then, I would like to highlight the importance of 1) taking into consideration the intensity of expectancy violations when studying their effects, and 2) introduce the idea that multiple psychological processes may compete to determine how people respond to expectancy violations. I conclude from the research reviewed above that it is critical to take into account individual differences that can shape people's motivation and capacity to respond to expectancy violations.

Hypothesis 7: Self-reinforcing cycles can perpetuate different cognitive responses.

ERT makes one final prediction——that there can be self-reinforcing psychological cycles wherein a psychological threat response is likely to fuel and reinforce defensive

processes, whereas a psychological challenge response is likely to fuel and reinforce open-minded processes (see "7. Self-reinforcing cycles" in Figure 7.2).

Hypothesis 7a: Psychological states can reinforce cognitive responses.

As an example, previous research has shown that negative emotions tend to be cognitively depleting (Morris & Feldman, 1996), and the negative affective states that characterise the threat response (such as fear and anxiety) have been associated with narrowed attention and thought-action repertoires (Fredrickson & Branigan, 2005). Depletion and narrowed attention, in turn, are likely to keep individuals in a threat / defensive mode (cf. research on the psychological effects of scarcity; Mullainathan & Shafir, 2013; Shah, Mullainathan, & Shafir, 2012). Because individuals in this mode will feel depleted, they are likely to continue drawing on existing schemata, be more ego-defensive, and less cognitively flexible. In contrast, the positive affective states that can characterise the challenge response, such as interest, curiosity, wonder, and awe, have been associated with a broadened scope of attention and wider thought-action repertoires (Fredrickson, 2001; Fredrickson & Branigan, 2005; Fredrickson & Levenson, 1998). This is likely to ease cognitive processing because more diverse information and associations will be accessible (Greenwald et al., 2002). ERT proposes that the experience of ease will further promote positive affective-motivational states and thus reinforce the link between the challenge response and open-minded cognitive processing (for a similar prediction, see Werner, Milyavskaya, Foxen-Craft, & Koestner, 2016).

Hypothesis 7b: Psychological traits can reinforce cognitive responses.

Moreover, ERT predicts that individuals with certain traits (e.g., those who have a high vs. low need for structure) are likely to inadvertently perpetuate their characteristic affective-motivational states (associated with indifference or threat or challenge) due to the cognitive and emotional habits that they have developed over time. Because individual differences such as need for structure are often deeply entrenched, indiscriminately relying on one's dominant response may be maladaptive and result in suboptimal outcomes. For example, individuals high in need for structure may have a tendency to respond rigidly and lack creativity, whereas individuals low in need for structure might display too much cognitive flexibility (Gocłowska et al., 2014). The latter tendency in particular could give rise to a "moral grey zone" that allows these individuals to justify a variety of situations in a self-serving manner (Gino & Ariely, 2012; Vincent, Emich & Goncalo, 2013). While I am not aware of any research testing whether indifference, threat, or challenge responses (as a result of individual differences) could self-perpetuate, I propose that this question be tested in future research.

7.4 ERT in the Context of Existing Models and Theories

ERT is a theory that is inspired by and, in part, built on the core elements of existing models and theories, in particular the meaning maintenance model (MMM;

Proulx & Inzlicht, 2012), the transactional model of stress and coping (Lazarus & Folkman, 1987), the CPAG model (Crisp & Turner, 2011), the diversifying experience model (DEM; Gocłowska, Damian, & Mor, 2017b), and the biopsychosocial (BPS) model of challenge and threat (Blascovich et al., 2003). However, it differs from these models because it provides a nuanced specification of the antecedents and consequences of responses to expectancy violations. In contrast to the MMM, ERT does not assume that expectancy violations are inherently aversive and represent threats. Instead, ERT postulates that a defensive response to expectancy violations constitutes only one possible psychological response to expectancy violations. Indifference/boredom or a psychological challenge response are alternative plausible psychological states following from expectancy violations. Unlike the MMM, ERT does not assume that proximal or distal reactions necessarily serve to reduce anxiety, but rather primarily serve to facilitate effective, adaptive behaviour (in line with Harmon-Jones, Amodio, & Harmon-Jones, 2009). As such, many expectancy violations are predicted to garner attention, rather than necessarily to arouse anxiety (which emotions are experienced in response to expectancy violations depends on individuals' motivation and capacity to deal with the unexpected information or event). Another difference between the MMM and ERT is that, while the MMM does not postulate the presence of any particular emotional states (other than anxiety) in response to expectancy violations, ERT makes directional predictions in line with the biopsychosocial (BPS) model of challenge and threat: A psychological threat response is predicted to involve negative emotional states (such as fear and stress) whereas a psychological challenge response is predicted to involve positive

emotional states (such as interest, curiosity, and awe).

While ERT is at least in part inspired by the transactional model of stress and coping, it is also different in a number of important ways. For example, the transactional model does not specify what happens if an expectancy violation is perceived as positive, or what happens if sufficient resources are available to deal with a negative expectancy violation. Notably, in subsequent work Tomaka et al. (1993) have supplemented this part of Lazarus' and Folkman's theory by establishing that a psychological challenge response can follow when sufficient resources are available to cope with the event. However, no comprehensive theory to date has tied together the psychological antecedents and consequences of expectancy violations (i.e., both negative and positive psychological consequences).

The CPAG model suggests that social and cultural diversity can promote open-minded thinking / cognitive flexibility provided that it challenges expectations, and provided that certain conditions are met (i.e., the perceiver is motivated and able to engage with the expectancy-violating information), which is consistent with ERT. Like the CPAG model, ERT also highlights the need to study the effects of expectancy violations over time as this would allow to assess their broader implications and persistence. However, ERT makes predictions beyond the realm of social and cultural diversity by accounting for any type of expectancy violation, be it social or non-social in nature. It also specifies two additional plausible responses to expectancy violations——indifference and threat——thus extending the scope of existing models. What is more, ERT suggests ways to help people to shift from viewing expectancy violations (including CSTs) as threatening, to viewing them as chal-

lenging, for example through reappraisal. ERT is domain-general and can thus be applied to any type of situation involving expectancy violations. Diversity science is one area that could be fruitfully informed by the theory.

The DEM suggests that the intensity of the expectancy violation (defined as the extent to which an experience requires cognitive adaptation) and adaptive resources (such as health, problem-solving skills, social support, material resources) are key moderators determining threat versus challenge responses following from expectancy violations, which is in line with ERT. However, ERT also suggests that it is important to consider differences in motivation (e.g., need for structure, need for cognition) as determinants of the psychological consequences expectancy violations can prompt. In line with this argument, Gocłowska, Damian, and Mor (2017b) briefly allude to epistemic needs (e.g., need for closure, intolerance of ambiguity) as potential constraints of when people actually use adaptive resources; however, this part of the model is not specified in more depth.

Lastly, ERT differs from the BPS model of challenge and threat in two important ways. First, a novel and central proposition of ERT is that expectancy violations can be antecedents of both threat and challenge responses. This extends the scope of the BPS model, which has historically been tested only in motivated performance situations and has not specified expectancy violations as precursors to threat or challenge. Specifically, Blascovich and Tomaka (1996) have postulated that goal-relevance is an important precondition of threat and challenge effects and argued that challenge and threat appraisals only occur in situations that are goal-relevant and evaluative. Here I argue that performance situations are a sufficient, but not

necessary, condition for threat versus challenge responses to occur. That is, threat and challenge responses are likely to follow from a variety of expectancy violations that are not goal-relevant or evaluative per se (e.g., see research on social schema violations and their effects on creative cognition, Gocłowska, Baas, Crisp, & De Dreu, 2014). ERT also extends the BPS model of threat and challenge by predicting that threat responses tend to be followed by cognitively defensive cognitive processing, whereas challenge responses tend to implicate cognitively expansive cognitive processing. To my knowledge, this prediction is novel and has only been partially or implicitly tested so far. In sum, ERT represents an attempt to incorporate a range of models and theories as part of one unifying theory.

7.5 Future Research

7.5.1 Reappraisal as a Strategy to Shift Responses

One important question for future research concerns the possibility to reappraise situations in order to shift from one state (e.g., a threat response) to another (e.g., a challenge response). From a psychological perspective, it is crucial for the mind to be flexible, and to be able to reappraise situations in the face of changing circumstances (Folkman, Lazarus, Dunkel-Schetter, DeLongis & Gruen, 1986; Gross, 1998; Park & Folkman, 1997). Some evidence indicates that shifting from a threat to a challenge response may indeed be possible. For example, Alter, Aronson, Darley, Rodriguez, and Ruble (2010) examined whether reframing threats as challenges could eliminate

the performance-inhibiting consequences of stereotype threat. The authors found that reframing math tests as tools to help improve abilities (termed the 'challenge' condition)——rather than diagnostic instruments that demonstrate ability (termed the 'threat' condition)——helped both school children and undergraduate students significantly improve performance in the challenge, but not in the threat condition. While emotional effects were not measured in these studies, it is very likely that participants in the challenge condition experienced significantly more positive emotions than participants in the threat condition (Blascovich et al., 2003; Tomaka et al., 1993). Alter et al. (2010) concluded that "(...) rather than merely dampening a threat, challenge-framing might recruit a motivational style that instantiates an adaptive stress-coping mechanism" (p. 170). I agree and propose that challenge framing may represent an adaptive stress-coping mechanism and hope to see more research in this area.

7.5.2 The Role of Self-awareness and Meta-cognitive Skills

One interesting question for future research is whether people are able to develop self-awareness or meta-cognitive knowledge that allows them to regularly recognise and switch between indifference, threat, and challenge responses to expectancy violations. Some evidence lends support to this possibility. Cross-cultural research suggests that people can develop meta-cognitive skills that help them to modify their behaviour in new contexts (the ability to "cross-culturally code-switch") in order to accommodate different cultural norms for appropriate behaviour (Molinsky,

2007). Moreover, if people develop a certain level of self-awareness and appropriate meta-cognitive skills that help them regulate their initial responses to expectancy violations, then this skill could have benefits for their health more generally. This is because some research suggests that the positive affective-motivational experiences implicated in the challenge response could speed the physical recovery from the cardiovascular repercussions of negative emotions (Fredrickson & Levenson, 1998). To ensure that individuals adapt effectively to unexpected life circumstances, either through personal introspection or with the help of psychological interventions (Walton, 2014), it is important that future research explores in more depth how people can develop self-awareness and self-regulatory skills when faced with expectancy violations.

7.5.3 Longitudinal Exposure to Expectancy Violations——What are the Adaptation Trajectories?

Because some types of expectancy violations happen on a repeated basis, an interesting question arises as to how people adapt to stimuli and events that violate their expectancies over time. For example, an individual who has lived in a small village all their life but has recently moved to a big city may be repeatedly exposed to new and unexpected forms of social and cultural diversity: He or she might meet a Muslim feminist, work with a hippie lawyer, or make friends with a disabled athlete. To my knowledge, little work has investigated how people cognitively and emotionally adapt to expectancy violations over time. One exception is Bettencourt

and Manning (2016), who examined how negative expectancy violations about the quality of life predicted emotionality in breast cancer survivors. The authors found that the more the patients perceived negative expectancy violations (i.e., agreed with statements such as "Compared to what I was expecting, the quality of my life at this time is much more negative") at earlier time points the more they experienced negative emotionality (e.g., depression, tension, anger, fatigue) at later time points. The authors concluded that life-threatening illness may engender chronic negative emotionality because individuals perceive that the quality of their lives is worse than expected. However, while this study makes a valuable contribution to the literature on expectancy violations by examining their longitudinal effects on emotional health, several questions remain: How exactly did the breast cancer survivors adapt to the expectancy violations over time? What were their adaptation trajectories, or the patterns of adaptation? Unfortunately, the index of negative emotionality used by Bettencourt and Manning subsumes emotions associated with a threat state (e.g., depression) or a challenge state (e.g., anger) into one composite measure. As a result, we cannot know whether the patients experienced a threat or challenge response. With ERT as the theoretical backdrop, future studies would be able to make predictions and test which emotional, motivational, and cognitive states expectancy violations trigger over time.

Finally, it is worth noting that emotional and cognitive adaptation to unexpected events can follow distinct trajectories. For example, it is possible that people adapt to (some types of) expectancy violations quickly. Consequently, these violations may have initial emotional or cognitive effects (e.g., surprise, curiosity), but

their effects may fade relatively fast and not prompt any major psychological downstream consequences. This is likely to be the case if the expectancy violations are low in intensity (e.g., a stranger was born in Japan, grew up in Mozambique, and has Irish nationality). In contrast, if the expectancy violations are high in intensity and potentially even repeated in nature (e.g., as might be expected in the case of an unexpected diagnosis of a chronic disease), then the adaptation trajectory is likely to differ. Taken together, it is important that future research tests these longitudinal hypotheses and develops interventions that not only alleviate the negative downstream consequences of expectancy violations, but also promote adaptive behaviour in response to unexpected events.

7.6 Conclusion

It is well established that humans have a basic psychological need to understand and predict events. However, often people's predictive ability is disrupted when they encounter surprising information or experience unexpected events. Existing models have predominantly focused on the negative and threat-related consequences of expectancy violations and the potential positive effects have been under-specified, despite empirical work demonstrating that such effects are possible. Expectancy Regulation Theory (ERT) is designed to explain how people respond to expectancy violations and proposes that, when people have sufficient motivation and capacity, expectancy violations can promote positive affective-motivational states (e.g., interest, curiosity, awe) and cognitively expansive processes that can help to find new

meanings (e.g., cognitive flexibility and creativity). The theory also points to individual differences that can influence people's motivation, capacity, and strategies helping them to change the way that they respond to stimuli and events that violate their expectations. My hope is that ERT will stimulate future theorising and research on expectancy violations and inform interventions that can help people to cope with——and even benefit from——unexpected events.

Chapter 8

General Discussion

This thesis contributes to the extant scientific literature by providing a programmatic examination of psychological adaptation to counter-stereotypical diversity, involving both empirical and theoretical development. The literature reviewed in Chapters 2, 3, and 4 yielded the directional prediction that exposure to counter-stereotypes (CSTs), under the right conditions, should enhance domain-general cognitive flexibility. The empirical work presented in Chapters 4, 5, and 6 provided evidence to suggest that simple exposure to CSTs may not be sufficient to enhance cognitive flexibility. However, the experiments reported in Chapter 5 indicate that people low (but not high) in need for cognition show improved cognitive reflection after exposure to CSTs. Chapter 6 provided evidence suggesting that repeated exposure to CSTs does not lead to changes in cognitive flexibility, intergroup bias, or concern for discrimination over time, although effects on surprise were found (as expected) and exploratory analyses revealed effects on epistemic unfreezing. Finally,

Chapter 7 presented a new theory (Expectancy Regulation Theory, ERT) aiming to specify the effects of unexpected information and events on affect, motivation, cognition, and behaviour. In this chapter, I summarise (a) the experimental results and (b) the hypotheses derived from ERT to discuss their theoretical and practical implications, limitations, and areas for future research.

8.1 Theoretical Background

> "And if you help everyone else in your worlds (...) to learn and understand about themselves and each other and the way everything works, and by showing them how to be kind instead of cruel, and patient instead of hasty, and cheerful instead of surly, and above all how to keep their minds *open and free and curious*... Then they will renew enough to replace what is lost (...)." (Philip Pullman, The Amber Spyglass, 2000, p. 496, emphasis added)

How do people psychologically adapt to diversity experiences that challenge stereotypes? Chapters 1 to 4 summarised the current state of the literature, irradiating the theoretical progress made and the empirical challenges ahead. Chapter 1 broadly delineated possible responses to growing social and cultural diversity comprising cultural separation, marginalisation, assimilation, and integration. In addition, the importance of managing rising diversity effectively was discussed. In Chapter 2, social diversity was conceptualised on a continuum, ranging from ancestral monocul-

tural societies, pluralistic societies, to modern socially complex societies. The focus of this thesis was placed on people's reactions and responses to the latter, rather complex and cross-cutting type of diversity. CSTs were conceptualised as special cases of expectancy violations and defined as combinations of social categories that are perceived as unusual and unexpected, thus bucking contemporary social norms. Two models——the continuum model of impression formation (Fiske & Neuberg, 1990) and the Categorization-Processing-Adaptation-Generalization (CPAG) model (Crisp & Turner, 2011)——provided the relevant theoretical backdrop in order to start answering the question posed at the beginning of this section. The continuum model of impression formation highlighted two fundamental types of cognitive processing that occur when people learn about social others: Category-based versus individuated processing, which represents a basic assumption made throughout this thesis. It can be argued that the CPAG model extended Fiske and Neuberg's model by postulating that certain types of diversity experiences——i.e., experiences that challenge stereotypes——can help individuals move away from category-based and towards individuated processing. According to the CPAG model, this process was predicted to have both immediate, and longitudinal effects (which were tested in Chapters 4, 5, and 6): 'One-shot' exposure to CSTs was predicted to activate a cognitively flexible mindset in the situation, whereas repeated exposure to CSTs was predicted to boost domain-general cognitive flexibility.

After having specified the hypotheses tested in this thesis, various empirical challenges for the study of counter-stereotypical diversity were discussed. Two paradigms that ought to be well suited for the study of diversity——the contact

and the exposure paradigm——were juxtaposed and their suitability for different settings discussed. The exposure paradigm (defined as learning information about an outgroup member or counter-stereotypical individual) was deemed more suitable for contexts that are characterised by high threat or high conflict or both because it would allow individuals to prepare for an intergroup encounter in a psychologically safe way, helping to pave the way for direct contact at a later stage.

In light of the prediction that counter-stereotypical experiences can affect broader cognitive functioning, two types of cognitive functioning were distilled that are particularly relevant to this thesis: (1) executive functioning (including cognitive reflection), and (2) cognitive flexibility and creativity. Whereas cognitive reflection was defined as the ability to resist reporting responses that first come to mind (see Chapters 4 and 5), cognitive flexibility was defined as the capacity to go beyond established and mentally accessible ways of thinking in favour of thinking differently from other people or differently from what is habitual (see Chapter 6).

Chapter 3 briefly recounted the history of the study of CSTs, and Chapter 4 summarised the current status of the literature regarding exposure to CSTs and its effects on cognitive reflection/flexibility. These latter two chapters provided an overview of empirical evidence testing the psychological effects of counter-stereotypical experiences and thus represented the empirical foundation on which the predictions in this thesis were built.

8.2 Summary of Findings

8.2.1 Chapter 4: Experiments 1–8

Chapter 4 presented eight experiments that were designed to conceptually replicate past findings on CSTs and cognitive reflection. Experiments 1 through 3 used an existing paradigm that involved imagined contact with a typical versus atypical Muslim. In Experiments 4 through 6 a new manipulation (imagined contact with a typical vs. atypical Chinese student) was developed to examine whether it could boost cognitive reflection. Experiments 7 and 8 employed a different, subtle manipulation of imagined counter-stereotypical contact (i.e., contact with a male vs. female mechanic). Despite strong manipulations of imagined counter-stereotypical contact, the results were highly inconsistent. A mini meta-analysis showed that on average, there was a trend opposite the predicted direction (d_+ = -0.13), which is a small effect (Cohen, 1988).

8.2.2 Chapter 5: Experiments 9–11

Chapter 5 examined whether people's motivation to cognitively engage with counter-stereotypical information (i.e., need for cognition, NFC) could influence the extent to which they engage in cognitive reflection. Two new CST interventions were developed and successfully pretested. Three experiments found consistent support for the hypothesis that exposure to CSTs promotes cognitive reflection for people low in NFC (d_+ = 0.36), but not for people high in NFC (d_+ = -0.21). The

meta-analytic average direct effect of the experimental versus control condition on cognitive reflection across Experiments 9–11 was statistically non-significant and very small ($d_+ = 0.08$). Moreover, the moderated mediation analysis showed that both participants low and high in NFC experienced more surprise, expectancy violation, and cognitive complexity, which explained improved performance for both groups. However, despite these two positive mediation effects, the net performance for people high in NFC was negative. As discussed in Section 5.9.1, it is possible that exposure to CSTs arouses different, potentially antagonistic component processes among people high in NFC——on the one hand, they may feel interested and intrigued, trying to make sense of the CSTs, but on the other hand they may get cognitively depleted or fatigued.

8.2.3 Chapter 6: Experiment 12

Chapter 6 presented and validated a new, well-controlled paradigm that can be used to experimentally study the effects of exposure to CSTs over time. This paradigm was used in a preregistered longitudinal experiment examining the effects of exposure to CSTs on cognitive flexibility, explicit bias, concern for discrimination, and "epistemic unfreezing" (a psychological process characterised by curiosity and open-mindedness). The results of the longitudinal experiment showed no statistically or practically significant effects for any of the confirmatory analyses (i.e. cognitive flexibility, explicit bias, concern for discrimination, and personal need for structure). However, effects were found for the exploratory analyses on epistemic

unfreezing, such that the CST intervention made participants more interested in and curious about the target individuals; however, this effect did not extend beyond the immediate context towards more general openness.

8.2.4 Chapter 7: Expectancy Regulation Theory

This chapter proposed a unifying theory aiming to integrate existing models and derive novel testable predictions answering the question: How do people psychologically adapt to unexpected information and events? Expectancy Regulation Theory (ERT) is based on the premise that people have a basic need to understand and predict events, which drives a cognitive monitoring system comprising a quick, heuristic system (System 1, the default) and a slow, reflective system (System 2). ERT predicts that whenever information or events violate expectancies (e.g., counter-stereotypical experiences), System 2 triggers a process determining how much motivation and capacity is available to deal with the expectancy violation. ERT further predicts that (1) insufficient motivation to deal with an expectancy violation (regardless of the person's capacity to do so) will typically result in indifference, (2) high motivation but low capacity to deal with an expectancy violation will result in a threat response, and (3) high motivation and high capacity will result in a challenge response. Importantly, ERT also predicts that a threat response will most likely lead to a defensive, close-minded cognitive mode, whereas a challenge response will likely give rise to an open-minded cognitive mode.

8.3 Theoretical Implications

The results of the 12 experiments were rather inconsistent, which poses a challenge to evaluating their theoretical implications. Perhaps the best starting point is to evaluate findings that were relatively unambiguous: The moderating role of NFC in the effect of counter-stereotypical experiences on cognitive reflection, and the effects of counter-stereotypical experiences on surprise and epistemic unfreezing. Regarding the former, the present work suggests that high levels of cognitive motivation may not be required in order for people to engage with CSTs, but instead a low level of cognitive motivation appears sufficient to enable cognitive reflection after exposure to CSTs. It seems that people with low cognitive motivation have more "head space"—that is, more potential to be cognitively stimulated—than those high in NFC, which the positive, small-to-medium-sized meta-analytic effect for people low in NFC suggests. However, it also appears that too much cognitive motivation can potentially backfire, which is indicated by the small negative meta-analytic effect for people high in NFC. Taken together, these findings represent a new boundary condition of the CPAG model. The model had postulated that people need to be motivated to engage in inconsistency resolution, yet the present findings suggest that a low level of cognitive motivation seems sufficient, whereas high levels of motivation may have ironic effects.

The latter set of findings, that is, the effects of counter-stereotypical experiences on surprise and epistemic unfreezing, corroborate past research (Hastie et al., 1990; Hutter et al., 2012; Kunda, Miller & Claire, 1990; Prati et al., 2015a) by

illuminating the motivational consequences for perceivers who are exposed to CSTs, which are underexplored. The finding that exposure to CSTs can promote interest and curiosity is particularly promising because (1) curiosity can be a very powerful driver of learning, motivation, and development (Litman et al., 2010; Silvia, 2008; von Stumm et al., 2011) and (2) it might open up a new area of research on the consequences of diversity for curiosity (assuming this exploratory result can be replicated and is not a false positive). In conjunction, the reported empirical findings suggest that motivational traits and states may play an important role in psychological adaptation to counter-stereotypical diversity. As such, they lend support to some of the hypotheses postulated as part of ERT (e.g., Hypothesis 4a: People need to be appropriately motivated in order to effectively deal with expectancy violations and Hypothesis 5c: Positive emotions and open-minded cognitive processing can follow as a result of expectancy violations).

Moreover, ERT itself has implications for theorising on psychological adaptation to counter-stereotypical diversity. First, building on existing models and research it predicts that expectancy violations may prompt a broader range of psychological consequences than previously assumed: Indifference versus threat versus challenge. Whereas the CPAG model assumed that one possible response is a cognitively flexible, open mindset (which is closely related to a challenge response), it did not specify what happens when people lack the motivation and ability to engage with expectancy-violating diversity experiences. In contrast, ERT suggests that other responses (indifference, threat) and corresponding cognitive states (e.g., defensiveness) are possible, depending individual and situational circumstances. This means that in

order to be able to predict how people psychologically adapt to counter-stereotypical experiences, one needs both theoretically consider, and empirically assess different types of stimuli and events (e.g., imagined vs. real counter-stereotypical contact) that can cause different kinds of responses (indifference vs. threat vs. challenge). In sum, ERT supplements the CPAG model by specifying when and how counter-stereotypical diversity might be met with resistance or cheer.

Other findings presented in this thesis were rather ambiguous and therefore difficult to interpret. Chapters 4, 5, and 6 do not lend support to the prediction that counter-stereotypical experiences directly boost cognitive reflection or flexibility. Of course, most of the experiments presented here were conceptual rather than direct replications, so it is possible that certain moderating variables——ones that were not accounted for in the reported experiments——may explain why the direct effect of counter-stereotypical experiences on cognitive flexibility was not replicated. For example, Vasiljevic and Crisp (2013) asked participants to generate counter-stereotypical exemplars themselves (i.e., write down five counter-stereotypical vs. stereotypical social category combinations), which is a task that presumably requires participants to be creative and prompts them to draw on exemplars they are familiar with in their personal lives. This experimental approach may render the counter-stereotypical exemplars more meaningful to participants, which could be a reason why they subsequently showed increased cognitive flexibility. However, the 12 experiments reported in this thesis nevertheless suggest that the direct effect of counter-stereotypical experiences on cognitive flexibility is likely to be more constrained than previously thought, such that a very specific set of situational

and/or individual conditions may be required that allows for the effect to arise, or that the presence of certain circumstances may easily undermine it (for example, fatigue, boredom, or motivated reasoning). As discussed in Section 4.7.1, this has implications for models and theories specifying the effects of expectancy-violating diversity experiences: In order to improve current theorising, it will be necessary to detail the optimising and boundary conditions of the effect of counter-stereotypical experiences on cognitive flexibility.

Further, Chapter 6 did not yield evidence for the hypothesis that counter-stereotypical experiences produce change in intergroup bias or concern for discrimination. Again, unspecified moderating variables could explain why this result deviates from past research. For instance, one possible moderator is the effect of time: Many studies examining the effects of CSTs on intergroup bias were conducted 10-30 years ago, and given that CSTs are notions that are appraised in the given moment depending on contemporary perceptions of stereotypes and social norms, it may be that the zeitgeist has changed so that today, people are less surprised or affected by CSTs than they used to be. Whereas the measures of surprise in the reported studies do suggest that participants felt more surprised by the counter-stereotypical than stereotypical target individuals, it may still be that the levels of surprise in response to CSTs have declined over time——CSTs may be seen as increasingly "normal" today, at least in liberal, individualistic societies. Unspecified moderators could similarly explain the non-significant finding for concern for discrimination. One difference between the longitudinal intervention by Forscher et al. (2017) and the present longitudinal experiment is that Forscher et al. recruited a US American

student sample, whereas the present work drew on a sample of the British general population. It is well documented that countries vary on different cultural dimensions (e.g., individualism vs. collectivism, uncertainty avoidance; Hofstede, 1983), which may have played a role in the effect of CSTs on concern for discrimination. It is important that future theories take into account the potential effects of cultural differences on psychological outcomes, so that diversity scientists can build a complete and holistic understanding of the psychological effects of expectancy-violating diversity experiences across cultures.

Finally, the results from Chapter 5 tentatively indicate that counter-stereotypical experiences may not simply trigger one or another psychological process, but potentially multiple competing processes. To be clear, the experiments do not provide direct evidence for this proposition. And yet, it is puzzling how a positive effect of exposure to CSTs on the composite mediator was found for people high in NFC, which resulted in improved cognitive reflection, but simultaneously overall cognitive reflection of people high in NFC decreased as a result of the intervention. As discussed in Section 5.9.1, it is important that future research considers the possibility of multiple competing processes in response to different types of diversity (for example, people feeling somewhat threatened and somewhat challenged by diversity at the same time).

8.4 Practical Implications

The results of Chapters 4, 5, and 6 suggest that more research needs to be done before one can confidently apply CST interventions in real life settings. Essentially, the experiments designed as part of this thesis were an attempt to achieve triangulation (Munafò & Smith, 2018): The goal was to validate past findings through different research methods and multiple lines of evidence. By all means, the failure to conceptually replicate a range of findings does not refute the original findings, but it does raise questions regarding how generalisable they are. If relatively small changes to the manipulations or measures caused the predicted effects to disappear, then the effects might be more elusive and not as strong as assumed. This, in turn, might mean that the reported experiments may have been under-powered and much larger sample sizes will be required in future research to detect potential effects.

One practical implication of the above is that it may be wise to focus on conducting direct replications before constructing elaborate conceptual replications because the latter may raise more questions than they can provide answers if the replication attempts are unsuccessful. After all, it is not possible to disentangle whether a conceptual replication failed because the original finding was spurious, or because of methodological differences between the original experiment and the conceptual replication. While it can be argued that direct replications may never be fully equivalent to original studies——because people and contexts change over time——it nonetheless seems important to replicate original findings as closely as possible, ideally in collaboration with the original authors (Brandt et al., 2014). This ap-

proach will allow scientists to draw conclusions about which effects are universal and durable.

Chapter 5 demonstrated that a simple "one size fits all" approach to employing CST interventions is overly simplistic and that individual differences need to be taken into account by both researchers and practitioners. Failing to do so could unintentionally give rise to adverse consequences for some individuals (e.g., people high in NFC) and it is therefore important from a research ethics perspective to minimise any damage that CST interventions may produce (e.g., by providing thorough debriefing).

The new paradigm developed in Chapter 6 can be used in future research to test predictions regarding the psychological effects of counter-stereotypical experiences (and to replicate previous findings). By design, this new intervention ensures that stereotype content is controlled for across conditions, therefore neatly manipulating exposure to CSTs. In addition, the paradigm lends itself for longitudinal research, which allows scientists to study the persistence of effects. Researchers could capitalise on this new paradigm in order to study psychological adaptation to counter-stereotypical experiences over time.

To summarise, it is critical that interventions are grounded in research and evaluated rigorously (Moss-Racusin et al., 2014), if they are to have the desired positive consequences in practice. More research, both in the laboratory and in the field, needs to be conducted before CST interventions can be applied in real life settings and confident predictions made about their likely effects, such as decreased intergroup bias. It appears that implementing diversity interventions that draw on CSTs,

in practical settings such as schools or companies, may be premature at this stage.

8.5 Limitations and Future Research

8.5.1 Construct Validity

The empirical work presented in this thesis is limited in how it operationalised counter-stereotypical experiences. The present research opted for a "top-down" or, theory-driven approach, wherein the experimenters developed CST interventions based on their conceptions thereof, and then measured participants' reactions. Future research may benefit from a "bottom-up" or, data-driven approach, which may allow to explore people's conceptualisations of and narratives about CSTs. In other words, a mixed methods approach may prove valuable (e.g., see Tashakkori & Teddlie, 2010), one that combines quantitative and qualitative methods (e.g., interventions that involve interview techniques), in order to better understand how people psychologically adapt to counter-stereotypical experiences.

It is also worth asking whether higher order constructs were accurately captured by the manipulations and measures used. Most of the constructs in this thesis were measured using established tasks and scales (e.g., cognitive reflection/flexibility; explicit bias; concern for discrimination; need for cognition; personal need for structure), but some were newly constructed and thus their validity and reliability is less certain (e.g., the composite mediator consisting of surprise, expectancy violation, and cognitive complexity in Chapter 5; epistemic unfreezing in Chapter 6). In all

experiments, the content of the stimulus materials was not pre-rated for valence (i.e., positivity vs. negativity), so there is uncertainty with respect to the potential effects of stimulus valence on the dependent variables. Thus, it will be important to validate the manipulations and assessment instruments in future research. For example, in order to ensure that the experimental versus control conditions have similar valence, one might ask an independent group of participants to pre-rate the stimuli in terms of valence. Or, one might examine the extent to which cognitive reflection and flexibility show discriminant versus convergent validity: Do they measure the same or different underlying constructs? There is some evidence that various measures that are pertinent to the topic at hand——such as NFC, epistemic curiosity, and openness for ideas——in fact, lack discriminant validity (Mussel, 2010). Further, the single-item measures of explicit bias in Chapter 6 were potentially limiting because they might not have captured different dimensions of bias (e.g., see Koch, Imhoff, Dotsch, Unkelbach & Alves, 2016). Thus, it may be worthwhile to capture the explicit bias construct using multiple items in future research. Besides, the epistemic unfreezing construct may potentially have different dimensions or facets, for example a domain-specific (e.g., "I would like to learn more about [name]") and domain-general component ("I'm generally feeling inquisitive, open-minded"). Future psychometric research is required to examine and establish construct validity of some of the constructs investigated in this thesis.

8.5.2 Internal Validity

Random assignment to conditions was successful in Experiments 9 through 12, as could be seen in the randomisation checks reported in Chapter 5, and in the lack of significant differences in the baseline measures of the longitudinal experiment in Chapter 6. It is currently not possible to ascertain whether random assignment to conditions was successful in Experiments 1 through 8. Because all experiments were double-blind (i.e., assignment to conditions was random and the experimenter did not know participants' conditions), it can be ruled out that the experimenter had any effects on the proceedings of the experiments. However, one limitation in Chapters 4 and 5 was that stereotype content slightly differed across conditions, which may potentially have confounded the effects in question. As discussed in Section 5.9.1, in Experiments 10 and 11 Mary was a schoolteacher in the control condition, but a political leader in the experimental condition. While both of these professional occupations require a certain level of expertise and qualifications, it is possible that one of the two——that is, Mary the political leader——was perceived as higher in status or power. This may (or may not) have caused changes in cognitive performance (e.g., by making participants feel more threatened in one condition than in the other), independent of the fact that the categories were cross-cutting and counter-stereotypical. The paradigm developed in Chapter 6 can be used to rule out this confound by holding the content of the intervention(s) constant across conditions.

Also note that most experiments presented in this thesis used stereotypical cat-

egorisation as a baseline condition, meaning that there were two conditions only——a stereotypical and a counter-stereotypical condition, but no condition devoid of categorisation. It could be argued that exposure to *stereotypes*, rather than CSTs, produced changes in some of the dependent variables, and that a neutral control condition was lacking. This may have prevented an evaluation of the directionality of the reported effects. For example, has exposure to CSTs led to epistemic unfreezing, or has exposure to stereotypes led to epistemic freezing? The experiments were purposefully designed with stereotypical categorisation as a baseline condition because it is generally accepted that stereotypic thinking is the default mode of person perception (e.g., Fiske & Neuberg, 1990). In support of this view, prior research found no differences between stereotypical and neutral / no categorisation conditions (e.g., Prati, Vasiljevic, et al., 2015b; Vasiljevic & Crisp, 2013). Given this evidence, I deemed it sufficient to contrast stereotypical versus counter-stereotypical conditions with each other. Plus, given the resource constraints, I aimed to maximise statistical power for the two existing conditions——having three conditions would have resulted in smaller samples per condition. Even so, in future research it may be worthwhile to include a neutral, no categorisation condition under certain circumstances in order to determine whether or not stereotypical categorisation may drive some, but not other effects.

8.5.3 External Validity

In this thesis, all participant samples were convenience samples, meaning that none of them were representative of the respective populations (UK and USA) in terms of gender, age, and other important demographic variables. As a result, it is not clear to what extent the reported findings are generalisable to the broader populations. Arguably, the participant samples that were recruited from Amazon's Mechanical Turk and Prolific Academic were relatively less biased than the samples from the two laboratory experiments. This is because the former two platforms allowed the general population to sign up and participate in research, whereas the latter almost exclusively drew on 1st year psychology undergraduate students only. Although these are limiting factors, the present research is consistent with the approach taken in published studies that reported significant findings. Future research would highly benefit from testing the predicted effects in nationally representative samples in order to increase the external validity of the research.

It is also important to reflect on what life stages in human development interventions may be best placed in. One limitation of this thesis is that it cannot speak to this question; ideally, future research would address it. Lai et al. (2016) speculated that perhaps diversity interventions, if they are to have significant impact in the long-term, might be better placed early in development. According to this sobering perspective, this is because change may be too difficult to achieve for adults. Research on implicit attitude change with children supports the idea that interventions targeting children's explicit and implicit biases can produce durable change

that lasts for years (e.g., Neto, Pinto, & Mullet, 2015). However, these programmes of research are in their early stages. Future research needs to test in a programmatic fashion at what point in human development diversity interventions may be best placed.

What is more, the construct of cognitive/integrative complexity deserves mention because it transcends basic definitions of cognitive functioning, and thus highlights the potentially complex nature of the processes involved when experiencing social and cultural diversity. Recall that in Chapter 2 this construct was defined as "the capacity and willingness to acknowledge the legitimacy of competing perspectives on the same issue (differentiation) and to forge conceptual links among these perspectives (integration)" (Tadmor, Tetlock & Peng, 2009, p. 105). Some research suggests that bicultural individuals are more cognitively complex than monocultural individuals (Benet-Martínez, Lee & Leu, 2006). Here, cognitive complexity was measured by asking participants to write ten statements to describe one of two cultures (US American vs. Chinese), which two coders rated on different complexity dimensions (e.g., whether the statement contained multiple perspectives, or by rating the overall complexity of the ideas or concepts contained in the statement). Further, research on multicultural engagement shows that immersion in multicultural environments not only promotes integrative complexity across domains (e.g., culture, work; Tadmor et al., 2009), but that it even predicts success in the job market via longitudinal increases in integrative complexity (Maddux, Bivolaru, Hafenbrack, Tadmor & Galinsky, 2014). It is important to consider these fruitful and informative areas in diversity science when developing one's theorising and planning new

empirical studies.

Finally, because the experiments presented in this thesis have only drawn on English-speaking samples, it is unclear how CSTs would be perceived in non-English-speaking cultures, and also non-Western cultures, and how counter-stereotypical experiences may impact emotion, cognition, motivation, and behaviour in such contexts. Research on social norms, which distinguishes between tight versus loose cultures, may render some interesting predictions for future research. Tight cultures have strong norms and a low tolerance of deviant behaviour (e.g., India, Pakistan, Malaysia), whereas loose cultures have weak norms and are relatively tolerant of deviant behaviour (e.g., Australia, Brazil, The Netherlands; Gelfand et al., 2011). Given that counter-stereotypical individuals often deviate from social norms, it may be predicted that people who hail from and are socialised in tight cultures may have different psychological adaptation trajectories when exposed to counter-stereotypical diversity than people from loose cultures. For example, it seems likely that on average, individuals from tight (vs. loose) cultures may show relatively high (vs. low) resistance to or dislike of CSTs. This is because chronic exposure to strong (vs. weak) social norms may suggest to individuals from tight (vs. loose) cultures that they have limited behavioural options, that their actions are being perpetually evaluated by society, and that they are likely to be punished for deviating or approving of deviancy. More broadly, research on the effects of social norms in different cultures underscores the importance of taking a more holistic view on diversity science, which considers multiple levels of analysis.

To summarise, the experiments presented in this thesis used a limited range

of CSTs to test the hypotheses and used only three measures of cognitive reflection/flexibility. Future research needs to test whether the effects discovered in the present research can be replicated with different manipulations and measures as well as in different contexts and cultures. In the following section, the sociocultural framework for diversity science is introduced, which will help future research formulate research questions at different levels of analysis.

8.5.4 A Sociocultural Framework for Diversity Science

As described in Chapters 1 and 2, the question of how people psychologically adapt to new forms of diversity has garnered significant research attention in the past three decades (Plaut, 2010b). To understand how different psychological realities are created and maintained, Markus and Kitayama (1994) suggested that it is imperative to take cultural factors and processes into account. In particular, the authors advise to take people's values, customs, norms, and societies' economic and socio-political circumstances into account when researching psychological processes underpinning people's sense of self and identity. Thus, Markus' and Kitayama's (1994) sociocultural approach to the study of the self complements the idea of the individualistic, independent nature of the self (typical in Western cultures) with a collectivistic, interdependent conception of the self (typical in East Asian cultures). Further, Markus and Kitayama argue that an "individualist ideal" pervasively influenced European-American social behaviour to the extent that even "(...) social psychologists, the very group committed to understanding the social nature of the

mind, approach the analysis of social behaviour with a distinctly asocial model of the self" (p. 568). In an attempt to correct for this imbalance, Markus and Kitayama proposed a framework that relates a set of macrolevel phenomena (e.g., cultural views of personhood) to a set of microlevel phenomena (e.g., cognitive, emotional, and motivational processes).

Plaut (2010b) built on this and proposed a sociocultural framework for diversity science. She argues that, to address the complex ethnic and racial issues of the 21st century, we require a diversity science, and to establish a diversity science, we need a sociocultural framework:

> "A diversity science will consider how people create, interpret, and maintain group differences among individuals, as well as the psychological and societal consequences of these distinctions. A diversity science will recognize that these significant social distinctions (...) are not simply natural, neutral, or abstract. Instead they are created and re-created in the process of everyday social interactions that are grounded in historically derived ideas and beliefs about difference and in a set of practices and institutions that reflect these ideas and beliefs. (...) [They] therefore shape psychological experience and behavior. According to this "sociocultural" framework, psychological experience and behavior, in turn, reinforce particular cultural and structural realities." (p. 77)

Plaut's sociocultural framework for diversity science highlights how individuals are both shaped by, and are architects of, their social worlds, much in line with

Markus' and Kitayama's model. Moving beyond an account that roots intergroup bias in people's individual psychologies, Plaut emphasises the need for a sociocultural analysis of diversity and its effects—that is, an examination of how daily experiences, practices and institutions, cultural ideas, and structural reality shape psychological structures and processes, and importantly, vice versa.

The sociocultural framework for diversity science is helpful in reflecting on the contributions of this thesis. First, it helps to organise different perspectives on diversity and intergroup relations by stressing the multi-level nature of phenomena related to diversity. It highlights that a psychological level of analysis presents an incomplete picture of diversity, and argues that scientists and practitioners would benefit from considering social and cultural levels of analysis (i.e., factors such as daily experiences, practices, institutions, and structural realities). Second, a sociocultural framework for diversity science enables social scientists to ask broader questions. For example, how do the institutions and practices that comprise education systems (e.g., teaching styles, organisational rules and structures) shape emotional, cognitive, and motivational processes and outcomes in students from diverse social and cultural backgrounds? If, for instance, Western teachers primarily value critical thinking and creativity (what is termed 'mind orientation'; Li, 2005), but are unaware that other cultures have contrasting conceptions of learning (e.g., emphasising the development of personal virtues such as respect, diligence, and perseverance, which has been called 'virtue orientation'; Li, 2005), then they may not be able to teach students from different cultures effectively, or cultural misunderstandings or even conflicts may arise (van Egmond, Kühnen & Li, 2013). Third

and last, the sociocultural framework locates the thesis in the context of relevant ideas and empirical work. To be precise, this thesis was predominantly concerned with psychological responses to diversity, which represent microlevel phenomena. It introduced new research and a new theory examining how exposure to counter-stereotypical diversity affects certain aspects of psychological functioning. In future research, it will be vital to take into account phenomena at different levels of analysis (as the sociocultural framework proposes), and to explore the interactions between different levels of analysis.

8.5.5 Reflections on Scientific Practices: Psychology's Renaissance

Two parallel developments characterise how psychology has recently developed as a scientific discipline. First, in an article published in 2011 and now cited over 2,500 times, Simmons, Nelson, and Simonsohn (2011) have raised concerns about researcher degrees of freedom and the extent to which they can increase the chances of false-positive findings (see also Gelman & Loken, 2014, for a discussion of the problems with data-dependent analysis, or the 'garden of forking paths'). More precisely, in the course of collecting and analysing data, researchers have the option to flexibly make a range of decisions——whether or not to collect more data, to exclude some observations, to include control variables, to combine or transform measures, or to report only a subset of experimental conditions. Simmons et al. (2011) discovered that these researcher degrees of freedom "made it unacceptably

easy (...) to accumulate (and report) statistically significant evidence for a false hypothesis" and to "present anything as [statistically] significant" (p. 1359).

Second, recent research has found that many published scientific findings might neither be reproducible (one is unable to generate the same quantitative results given the same data as input), nor reliable (one is not able to replicate findings by collecting new data). For example, the Open Science Collaboration (2015)——consisting of research teams from around the world——conducted replications of 100 classic studies from psychology using high-powered designs and original materials where possible. In short, the authors were only able to replicate approximately 40% of the original scientific results, and the perspectives on this finding vary drastically. In an on-going debate, researchers on one end of the spectrum worry about the health of the field, wondering whether psychological science is truly self-correcting (Open Science Collaboration, 2015; Shrout & Rodgers, 2018), whereas researchers on the other end of the spectrum see no reason to be alarmed and deem the conclusions reached by the Open Science Collaboration pessimistic (Gilbert, King, Pettigrew & Wilson, 2016).

Importantly, as an antidote to the described problems, researchers have started to discuss, research, and adopt open science practices, which some have called the 'open science movement' (Gilbert & Corker, 2017). Open science practices help make scientific processes and results more transparent and reproducible by making research materials, data, and analysis code openly available. Thus, by allowing researchers to independently validate findings, they allow the field to grow and improve faster than it would otherwise. What is more, some have suggested that 'the

preregistration revolution' is under way (Nosek, Ebersole, DeHaven & Mellor, 2018). Preregistration helps distinguish analyses and outcomes that result from prediction (i.e., confirmatory, hypothesis-testing research) from those that result from postdiction (i.e., exploratory, hypothesis-generating research)(see also Kerr, 1998). Nosek et al. (2018) explain why this distinction matters: (1) failing to appreciate it can "lead to overconfidence in post hoc explanations (postdictions)", which inflate "the likelihood of believing that there is evidence for a finding when there is not" due to hindsight bias (also known as the 'I-knew-it-all-along effect'), and (2) presenting postdictions as predictions can make findings more attractive and publishable by falsely reducing uncertainty, which ultimately decreases reproducibility (p. 1). If the goal of science is to discern truth and develop knowledge, then all of the above concerns have to be taken very seriously.

Thankfully, recent years have seen a surge of new platforms and tools that are developed to help address problems like researcher degrees of freedom and to facilitate open science practices. The Open Science Framework (OSF; www.osf.io) is an open source platform that facilitates open collaboration in scientific research. With recent projects like the preregistration challenge (www.cos.io/prereg), the OSF attempts to incentivise and increase the adoption of preregistration protocols in the scholarly community, which will ultimately help enhance credibility of scientific findings. Other developments include the adoption of badges by scientific journals such as *Psychological Science*, which signal open science practices (e.g., open materials badge; open data badge). One study revealed that badges are indeed an effective, and at the same time simple and low-cost, method for increasing transparency:

Before badges were introduced in January 2014, less than 3% of Psychological Science articles reported open data; this number rose to 39% in the first half of 2015, which is a substantial increase. Taken together, it has been argued that this period of methodological reflection has dramatically improved experimental psychologists' scientific practices, which has been termed 'Psychology's Renaissance' (Nelson, Simmons & Simonsohn, 2018).

In light of the above, it is worthwhile to take a self-critical look at the work presented here. This thesis has adhered to open science standards some, but not all of the time, which presents a limitation of the present work. For example, Experiment 12 is the only one, out of 12 experiments, that was preregistered, and it can therefore not be ruled out that researcher degrees of freedom influenced the data collection and analyses in the remaining eleven experiments. Further, because the early experiments (Experiments 1–8) were analysed using point-and-click-tools (i.e., SPSS) rather than programming languages, there is no analysis code that can be shared. While this certainly does not prevent researchers from reproducing the findings, it does put obstacles in the way of reproducibility because it makes it difficult to retrace what exact assumptions and parameters the analyses were based on. For future research testing the effects of diversity, it is recommended that researchers (1) openly share their study materials, data, and analysis code in order to facilitate scientific progress, and (2) preregister their hypotheses in order to reduce researcher degrees of freedom and improve the credibility of scientific findings. This will hopefully make research on psychological adaptation to diversity specifically, and psychological science more generally, more reliable and valid.

8.6 Conclusions

In conclusion, this thesis aimed to make both empirical and theoretical contributions to the study of psychological adaptation to counter-stereotypical diversity. One primary prediction was derived from the literature, which is that exposure to CSTs can boost cognitive flexibility, and tested across 12 experiments (Chapters 4, 5, and 6). Various secondary predictions were tested as well, for example the role of NFC in the effects of exposure to CSTs (Chapter 5), and the longitudinal effects of counter-stereotypical experiences (Chapter 6). CSTs were conceptualised as a special case of expectancy violations, and Chapter 7 theorised that they can be followed by three types of responses: Indifference versus threat (and defensiveness) versus challenge (and open-mindedness). This final chapter critically discussed various theoretical and practical implications of the present work as well as limitations and promising avenues for future research. Hopefully, this thesis could demonstrate that the study of people's psychological experiences of (counter-stereotypical) diversity is a worthwhile pursuit that has the potential to improve both individual and societal welfare.

Appendix A

Experiments 1–3: Heuristics and Biases Problems

Problem 1

A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50 percent of all babies are boys. However, the exact percentage varies from day to day. Some-times it may be higher than 50 percent, sometimes lower. For a period of 1 year, each hospital recorded the days on which more than 60 percent of the babies born were boys. Which hospital do you think recorded more such days? Circle one:

(a) The larger hospital

(b) The smaller hospital

(c) About the same (that is, within 5 percent of each other)

Problem 2

A game of squash can be played either to 9 or to 15 points. Holding all other rules of the game constant, if A is a better player than B, which scoring system will give player A a better chance of winning?

(a) 9 points

(b) 15 points

Problem 3

After the first 2 weeks of the major league baseball season, newspapers begin to print the top 10 batting averages. Typically, after 2 weeks, the leading batter often has an average of about .450. However, no batter in major league history has ever averaged .450 at the end of the season. Why do you think this is? Circle one:

(a) When a batter is known to be hitting for a high average, pitchers bear down more when they pitch to him.

(b) Pitchers tend to get better over the course of a season, as they get more in shape. As pitchers improve, they are more likely to strike out batters, so batters' averages go down.

(c) A player's high average at the beginning of the season may be just luck. The longer season provides a more realistic test of a batter's skill.

(d) A batter who has such a hot streak at the beginning of the season is under a lot of stress to maintain his performance record. Such stress adversely affects his

playing.

(e) When a batter is known to be hitting for a high average, he stops getting good pitches to hit. Instead, pitchers "play the corners" of the plate because they don't mind walking him.

Problem 4

When playing slot machines, people win something about 1 in every 10 times. Julie, however, has just won on her first three plays. What are her chances of winning the next time she plays?

Problem 5

A doctor had been working on a cure for a mysterious disease. Finally, he created a drug that he thinks will cure people of the disease. Before he can begin to use it regularly, he has to test the drug. He selected 300 people who had the disease and gave them the drug to see what happened. He selected 100 people who had the disease and did not give them the drug in order to see what happened. The table below indicates what the outcome of the experiment was:

Was the treatment positively or negatively associated with the cure for this disease? Please indicate your judgment on the following scale from –10 (strong negative association) to +10 (strong positive association) by circling one of the numbers:

-10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 +1 +2 +3 +4 +5 +6 +7 +8 +9 +10

Problem 6

The city of Middleopolis has had an unpopular police chief for a year and a half. He is a political appointee who is a crony of the mayor, and he had little previous experience in police administration when he was appointed. The mayor has recently defended the chief in public, announcing that in the time since he took office, crime rates decreased by 12%. Which of the following pieces of evidence would most deflate the mayor's claim that his chief is competent?

(a) The crime rates of the two cities closest to Middleopolis in location and size have decreased by 18% in the same period.

(b) An independent survey of the citizens of Middleopolis shows that 40% more crime is reported by respondents in the survey than is reported in police records.

(c) Common sense indicates that there is little a police chief can do to lower crime rates. These are for the most part due to social and economic conditions beyond the control of officials.

(d) The police chief has been discovered to have business contacts with people who are known to be involved in organized crime.

Problem 7

Assume that you are presented with two trays of black and white marbles: a large tray that contains 100 marbles and a small tray that contains 10 marbles. The marbles are spread in a single layer on each tray. You must draw out one marble (without peeking, of course) from either tray. If you draw a black marble, you win £2. Consider a condition in which the small tray contains 1 black marble and 9

white marbles, and the large tray contains 8 black marbles and 92 white marbles. From which tray would you prefer to select a marble in a real situation?

(a) Large tray

(b) Small tray

Problem 8

1. You are staying in a hotel room on vacation. You paid £6.95 to see a movie on pay TV. After 5 minutes you are bored and the movie seems pretty bad. Would you continue to watch the movie or not?

(a) Continue to watch

(b) Turn it off

2. You are staying in a hotel room on vacation. You turn on the TV and there is a movie on. After 5 minutes you are bored and the movie seems pretty bad. Would you continue to watch the movie or not?

(a) Continue to watch

(b) Turn it off

Problem 9

Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which of the following options is likely? Circle one:

(a) Linda is a bank teller.

(b) Linda is a bank teller and is active in the feminist movement.

Problem 10

The Caldwells had long ago decided that when it was time to replace their car they would get what they called "one of those solid, safety-conscious, built-to-last Swedish cars"-either a Volvo or a Saab. As luck would have it, their old car gave up the ghost on the last day of the closeout sale for the model year both for the Volvo and for the Saab. The model year was changing for both cars and the dollar had recently dropped substantially against European currencies; therefore, if they waited to buy either a Volvo or a Saab, it would cost them substantially more-about £1200. They quickly got out their Consumer Reports where they found that the consensus of the experts was that both cars were very sound mechanically, although the Volvo was felt to be slightly superior on some dimensions. They also found that the readers of Consumer Reports who owned a Volvo reported having somewhat fewer mechanical problems than owners of Saabs. They were about to go and strike a bargain with the Volvo dealer when Mr. Caldwell remembered that they had two friends who owned a Saab and one who owned a Volvo. Mr. Caldwell called up the friends. Both Saab owners reported having had a few mechanical problems but nothing major. The Volvo owner exploded when asked how he liked his car. "First that fancy fuel injection computer thing went out: £250. Next I started having trouble with the rear end. Had to replace it. Then the transmission and the clutch. I finally sold it after 3 years for junk." Given that the Caldwells are going to buy

either a Volvo or a Saab today, in order to save £1200, which do you think they should buy? Why?

Appendix B

Experiments 4–6: New Manipulations

Traditional Muslim

Imagine yourself traveling to Turkey for the first time. During your stay you get the chance to meet a Turkish woman with a decent command of English. Based on her clothing and on the conversation with her, you understand that she is a devout Muslim. That is, she dresses in a traditional way, avoids alcohol, reads the Koran and prays five times a day. Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner.

Muslim feminist

Imagine yourself traveling to Turkey for the first time. During your stay you get the chance to meet a Turkish woman with a decent command of English. Based on her clothing and on the conversation with her, you understand that she is a devout Muslim. That is, she dresses in a traditional way, avoids alcohol,

reads the Koran and prays five times a day. At the same time, she questions standard gender roles, advocates equal rights for women and is regularly engaged in feminist campaigns — she is a Muslim feminist. Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner.

Typical Chinese student

Imagine yourself traveling to China for the first time. During your stay you get the chance to meet a Chinese student with a decent command of English. After some time conversing with the student, you realise that this student is harmony-oriented, compliant and diligent. Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner.

Atypical Chinese student

Imagine yourself traveling to China for the first time. During your stay you get the chance to meet a Chinese student with a decent command of English. After some time conversing with the student, you realise that this student is rebellious, individualistic and quite easy-going. Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner.

Heterosexual soldier

Imagine yourself traveling to Russia for the first time. During your stay you happen to meet a young soldier with a decent command of English and get the chance to talk to him for a while. From the conversation you find out that he holds

the army in high esteem, that he is proud to serve his nation, and that he considers it important to abide by orders of authority figures. Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner.

Homosexual soldier

Imagine yourself traveling to Russia for the first time. During your stay you happen to meet a young soldier with a decent command of English and get the chance to talk to him for a while. From the conversation you find out that he holds the army in high esteem, that he is proud to serve his nation, and that he considers it important to abide by orders of authority figures. You then also find out that he is homosexual, which is unknown to his comrades. Imagine that you enjoy the interaction, and try to visualise the life and experiences of your conversation partner.

Appendix C

Experiment 9: Attention Check

Please tell us whether you would like to continue with this study. In fact, we are interested in determining if people read instructions. You must answer this question correctly in order to participate in this study. Please select the 'leave the study' option and then click on the arrow below to continue.

Would you like to continue with this study?

o Continue with the study

o Leave the study

Appendix D

Experiment 11: Attention Check

Please indicate your agreement with the question and statements below.

[Scale from 0 (not at all) to 100 (very much)]

(. . .) "Please move the slider to number eighty." (. . .)

Appendix E

Experiments 9–11: Composite Mediator

Items constituting the new composite mediator variable Measured on a scale from 0 (not at all) to 100 (very much)

1. **Surprise**: "To what extent do you feel surprised?"

2. **Perceived expectancy violation**:

"(name) is a typical (insert group membership)" (reverse-coded);

"Reading about (name) challenged some of my beliefs";

"There isn't anything puzzling about (name)'s life", reverse-coded;

"Imagining (name)'s life made me think 'outside the box'"

3. **Cognitive complexity**: "Forming an impression about (name) was a complex task."

Appendix F

Experiment 12: Samples of Imagery Descriptions

Descriptions provided by participants in the **intervention group:**

"I am intrigued by Rebecca as she has views which I do not generally see both held by the same person. I imagine her to be quite strong willed and forthright in her opinion. She does not just follow the crowd."
"Not believing in life after death seems strange if she is committed to praying five times a day. I'm curious as to why."

"I think Ahmed is a young, moderate Muslim who feels at home with Western democracy and has a relatively liberal world view. I think he probably feels incredibly uncomfortable with the more conservative strains of Islam and strives for a more progressive outlook within his faith."

"Clearly and regrettably relatively few car mechanics are female. One wonders therefore what institutionalised resentment is faced by those who are, and how their situation varies in different cultures and circumstances."

Descriptions provided by participants in the **control group:**

"Abdul is faithful. He works in a shop and came to this country when he was 14 years old with his parents who were seeking asylum. He reads the Koran every day and thinks that too many people are ungodly."

"I think Mary is warm and bubbly. She is easy going and gets on well with most people. I think Mary is funny and talkative with a big personality. I think Mary might be a little over-weight as she likes to eat and doesn't have time to exercise."

"Ahmed is a Muslim and was born into that religion. He has fasted since he was 10 years old. He is very devout, but also enjoys the communal aspects of fasting and the fact that there is a huge party once the fast is over. he works in a shop and only really socialises with other Muslims."

"I imagine Paul is a very typical working-class male. I don't think he cares too much education and gets all of his information from tabloid newspapers (such as the Sun). I think he enjoys drinking alcohol and he spends lots of time in the pub with friends who all have nicknames closely related to their surname."

Appendix G

Experiment 12: Scoring Instructions

1.) Study

Participants were asked to take part in a multi-part study over the course of 14 days. Every two days, participants were instructed to come up with names for new products or items (presented in a random order). In a separate excel sheet, you will find a list of participants' responses to these seven tasks:

(1) We would like you to think of various names for a new kind of PASTA, for example "fussilini", "paragoni", "malwini".

(2) We would like you to think of various new names for recently discovered PLANETS, for example "verunus", "arctanus", "tronus".

(3) We would like you to think of various names for newly discovered FLOWERS, for example "lunia", "fridia", "ezilia".

(4) We would like you to think of various names for new SOFTWARE COMPANIES, for example "Triddle", "Wubble", "Kimple".

(5) We would like you to think of various names for new AIRLINES, for example "Jimair", "Greenair", "Scanair".

(6) We would like you to think of various new names for GREEK ISLANDS, for example "Mianos", "Nikonos", "Presos".

(7) We would like you to think of various names for new COCKTAILS, for example "domicita", "hawaiana", "passilada".

2.) Scoring instructions

Your task is to score the responses in the excel sheet in terms of how diverse they are. In other words, you will judge for each participant how different/dissimilar the responses (in each set of responses) are. This is what the excel sheet should look like: "responseid" in column A, "PROLIFICPID" in column B, "flexnumber" in column C, and "flexexample" in column D, "flexresponse" in column E, and "flexrating in column F.

Make sure that you click on "Wrap Text" for column E so that the responses of each participant are visible at once. Before rating each set of responses, please have a read through the different versions of the task (see above) in order to familiarize yourself with the examples given as part of the instructions (e.g., "lunia", "fridia", "ezilia"). Then, please have a read through all of the responses that participants gave. That is, just look through the responses to get a sense of what kinds of responses participants gave. What are common answers? What are dissimilar and unique answers? Responses can be unique and dissimilar **if they *sound* different, have different *length*, relate to diverse *categories* of words, *start* with**

different syllables, or *end* with different syllables. Just briefly read through all responses and try to get an idea of the responses that participants gave.

Next, your task will be to rate the responses (within each set of responses) on a scale from 1 (not at all diverse) to 5 (very diverse). In assigning your ratings, try to take into account whether participants' responses are different from the provided examples (i.e., column D) or not.

If a set of responses contains many responses that are very diverse and dissimilar from each other, then it should get a rating of 4 or 5 (e.g. "xexxe, tanolu, ateni". If a set of responses contains many responses that are quite similar to each other (e.g., "lulia, lelia, lilia"), then it should get a rating of 1 or 2. If a set of responses contains some responses that are very similar and some responses that are very dissimilar, then it should get a rating of 3.

In addition, responses that are very similar *to the provided examples* should get a rather low rating. For example, if the examples are "lunia, fridia, ezilia" and participants' responses are "lulia, fredia, esilia", then the rating should be a 1 or 2. In contrast, responses that noticeably diverge from the provided examples should get a rather high rating. For example, if the examples are "lunia , fridia, ezilia" and participants' responses are "xexxe, tanolu, ateni", then the rating should be a 4 or 5.

If any responses are missing completely, then treat these simply as missing responses (i.e., leave them blank), rather than assigning a rating. If any of the responses represent existing names (for example, spaghetti, penne), then simply ignore these and evaluate all remaining responses that actually constitute new names.

Lastly, if participants re-use a lot of syllables (for example, airgo; goair; flyfar; farfly; travelfly; greenfly; greenair; flyair; airtime; flytime; flywithme; flywithus; flyhigh; highfly; up high fly; highair; jumboair; airtime), then make sure not to "punish" them if their responses are otherwise diverse. For example, if the participant in the last example had only written 50% of the responses, would they have received a reasonably high score? If yes, then all additional responses should not detract from the high score.

While keeping the above in mind, now please assign a single rating (either 1, 2, 3, 4 or 5) to each set of responses in column F named (flexrating). Remember, your rating should depend on how diverse/different the responses are (both as compared to each other and as compared to the examples).

References

Aberson, C. L. (2015). Positive intergroup contact, negative intergroup contact, and threat as predictors of cognitive and affective dimensions of prejudice. *Group Processes and Intergroup Relations*, *18*(6), 743–760. doi:10.1177/1368430214556699

Abrams, D. & Hogg, M. A. (1988). Comments on the motivational status of self-esteem in social identity and intergroup discrimination. *European Journal of Social Psychology*, *18*(4), 317–334. doi:10.1002/ejsp.2420180403

Aiken, L. S. & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage Publications.

Ainsworth, M. S. (1979). Infant–mother attachment. *American psychologist*, *34*(10), 932–937. doi:10.1037/0003-066X.34.10.932

Alesina, A. F. & La Ferrara, E. (2002). Who trusts others? *Journal of Public Economics*, *85*(2), 207–234. doi:10.1016/S0047-2727(01)00084-6

Allen, T. J., Sherman, J. W., Conrey, F. R. & Stroessner, S. J. (2009). Stereotype strength and attentional bias: Preference for confirming versus disconfirming

information depends on processing capacity. *Journal of Experimental Social Psychology*, *45*(5), 1081–1087. doi:10.1016/j.jesp.2009.06.002

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

Allport, G. W. (1968). The historical background of modern social psychology. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (2nd ed., p. 80). Cambridge, MA: Addison-Wesley.

Allport, G. W. (1979). *The nature of prejudice*. Basic Books.

Alter, A. L., Aronson, J., Darley, J. M., Rodriguez, C. & Ruble, D. N. (2010). Rising to the threat: Reducing stereotype threat by reframing the threat as a challenge. *Journal of Experimental Social Psychology*, *46*(1), 166–171. doi:10.1016/j.jesp.2009.09.014

Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology*, *45*(2), 357–376. doi:10.1037/0022-3514.45.2.357

Anchalee, I. (2016). #ILookLikeAnEngineer: One year later. Retrieved April 20, 2018, from https://medium.com/startup-grind/ilooklikeanengineer-one-year-later-b599e0cae817

Anderson, C. J. & Paskeviciute, A. (2006). How ethnic and linguistic heterogeneity influence the prospects for civil society: A comparative study of citizenship behavior. *The Journal of Politics*, *68*(4), 783–802. doi:10.1111/j.1468-2508.2006.00470.x

Anti-Prejudice Consortium. (2016). Power Over Prejudice. Retrieved April 21, 2018, from http://poweroverprejudice.org/

Aristotle. (350BC). Metaphysics. Retrieved November 27, 2017, from http://classics.mit.edu/Aristotle/metaphysics.html

Aubrey, J. S. & Harrison, K. (2004). The gender-role content of children's favorite television programs and its links to their gender-related perceptions. *Media Psychology*, *6*(2), 111–146. doi:https://doi.org/10.1207/s1532785xmep0602_1

Baas, M., De Dreu, C. K. W. & Nijstad, B. A. (2008). A meta-analysis of 25 years of mood-creativity research: Hedonic tone, activation, or regulatory focus? *Psychological Bulletin*, *134*(6), 779–806. doi:10.1037/a0012815

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, *49*(A), 5–28.

Baddeley, A. D. & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. doi:10.1016/S0079-7421(08)60452-1

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*(2), 122–147.

Barlow, F. K., Paolini, S., Pedersen, A., Hornsey, M. J., Radke, H. R. M., Harwood, J., ... Sibley, C. G. (2012). The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice. *Personality and Social Psychology Bulletin*, *38*(12), 1629–1643. doi:10.1177/0146167212457953

Bartholow, B. D., Fabiani, M., Gratton, G. & Bettencourt, B. A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science*, *12*(3), 197–204. doi:10.1111/1467-9280.00336

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01

Batson, C. D. (1975). Rational processing or rationalization? The effect of disconfirming information on a stated religious belief. *Journal of Personality and Social Psychology*, *32*(1), 176–184. doi:10.1037/h0076771

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

Baumeister, R. F. & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529. doi:10.1037/0033-2909.117.3.497

Benet-Martínez, V., Lee, F. & Leu, J. (2006). Biculturalism and cognitive complexity: Expertise in cultural representations. *Journal of Cross-Cultural Psychology*, *37*(4), 386–407. doi:10.1177/0022022106288476

Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology*, *45*(3), 180–191. doi:10.1111/j.2044-8295.1954.tb01243.x

Berry, J. W. (1970). Marginality, stress and ethnic identification in an acculturated aboriginal community. *Journal of Cross-Cultural Psychology*, *1*(3), 239–252. doi:10.1177/135910457000100303

Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied Psychology*, *46*(1), 5–68. doi:10.1111/j.1464-0597.1997.tb01087.x

Berry, J. W. (2001). A psychology of immigration. *Journal of Social Issues*, *57*(3), 615–631. doi:10.1111/0022-4537.00231

Bettencourt, B. A. & Manning, M. (2016). Negatively valenced expectancy violation predicts emotionality: A longitudinal analysis. *Emotion*, *16*(6), 787–791. doi:10.1037/emo0000152

Blair, I. V. & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, *70*(6), 1142–1163. doi:10.1037/0022-3514.70.6.1142

Blair, I. V., Ma, J. E. & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of personality and social psychology*, *81*(5), 828–841. doi:10.1037/0022-3514.81.5.828

Blascovich, J., Mendes, W. B., Hunter, S. B., Lickel, B. & Kowai-Bell, N. (2001). Perceiver threat in social interactions with stigmatized others. *Journal of Personality and Social Psychology*, *80*(2), 253–267. doi:10.1037/0022-3514.80.2.253

Blascovich, J., Mendes, W. B., Tomaka, J., Salomon, K. & Seery, M. (2003). The robust nature of the biopsychosocial model challenge and threat: A reply to Wright and Kirby. *Personality and Social Psychology Review*, *7*(3), 234–43. doi:10.1207/S15327957PSPR0703_03

Blascovich, J. & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. *Advances in Experimental Social Psychology*, *28*, 1–51. doi:10.1016/S0065-2601(08)60235-X

Boot, N., Baas, M., Mühlfeld, E., De Dreu, C. K. W. & Van Gaal, S. (2017). Widespread neural oscillations in the delta band dissociate rule convergence from

rule divergence during creative idea generation. *Neuropsychologia*, *104*, 8–17. doi:10.1016/j.neuropsychologia.2017.07.033

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. New York: Basic Books.

Brambilla, M., Ravenna, M. & Hewstone, M. (2012). Changing stereotype content through mental imagery: Imagining intergroup contact promotes stereotype change. *Group Processes and Intergroup Relations*, *15*(3), 305–315. doi:10.1177/1368430211427574

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., ... van't Veer, A. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, *50*, 217–224. doi:10.1016/J.JESP.2013.10.005

Brehm, J. W. (1966). *A theory of psychological reactance*. San Diego, CA: Academic Press. doi:10.1002/hrdq.20027

Brigham, J. C. (1993). College students' racial attitudes. *Journal of Applied Social Psychology*, *23*(23), 1933–1967. doi:10.1111/j.1559-1816.1993.tb01074.x

Brown, R. & Hewstone, M. (2005). An integrative theory of intergroup contact. *Advances in Experimental Social Psychology*, *37*, 255–343. doi:10.1016/S0065-2601(05)37005-5

Bruner, J. S. & Postman, L. (1949). On the perception of incongruity: A paradigm. *Journal of Personality*, *18*(2), 206–223. doi:10.1111/j.1467-6494.1949.tb01241.x

Buhrmester, M., Kwang, T. & Gosling, S. D. (2011). Amazon's Mechanical Turk. *Perspectives on Psychological Science*, *6*(1), 3–5. doi:10 . 1177 / 1745691610393980

Burgoon, J. K. (1993). Interpersonal expectations, expectancy violations, and emotional communication. *Journal of Language and Social Psychology*, *12*(1-2), 30–48. doi:https://doi.org/10.1177/0261927X93121003

Cacioppo, J. T., Petty, R. E., Feinstein, J. A. & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*(2), 197–253. doi:10.1037/0033-2909.119.2.197

Cacioppo, J. T., Petty, R. E. & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of personality assessment*, *48*(3), 306–307. doi:10.1207/s15327752jpa4803_13

Cameron, L., Rutland, A., Turner, R. N., Holman-Nicolas, R. & Powell, C. (2011). Changing attitudes with a little imagination: Imagined contact effects on young children's intergroup bias. *Annals of Psychology*, *27*(3), 708–717.

Campbell, D. T. (1965). Ethnocentric and other altruistic motives. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol. 13, pp. 283–311). Lincoln: University of Nebraska Press. doi:BF683N4V.13

Camras, L. A., Meng, Z., Ujiie, T., Dharamsi, S., Miyake, K., Oster, H., ... Campos, J. (2002). Observing emotion in infants: Facial expression, body behavior, and rater judgments of responses to an expectancy-violating event. *Emotion*, *2*(2), 179–193. doi:10.1037//1528-3542.2.2.179

Caporael, L. R. (1997). The evolution of truly social cognition: The core configurations model. *Personality and Social Psychology Review*, *1*(4), 276–298. doi:https://doi.org/10.1207/s15327957pspr0104_1

Cheng, C. Y., Leung, A. K. Y. & Wu, T. Y. (2011). Going beyond the multicultural experience-creativity link: The mediating role of emotions. *Journal of Social Issues*, *67*(4), 806–824. doi:DOI10.1111/j.1540-4560.2011.01729.x

Christ, O., Schmid, K., Lolliot, S., Swart, H., Stolle, D., Tausch, N., ... Hewstone, M. (2014). Contextual effect of positive intergroup contact on outgroup prejudice. *Proceedings of the National Academy of Sciences*, *111*(11), 3996–4000. doi:10.1073/pnas.1320901111

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. doi:10.1017/S0140525X12000477

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen-Chen, S., Halperin, E., Crisp, R. J. & Gross, J. J. (2014). Hope in the Middle East: Malleability beliefs, hope, and the willingness to compromise for peace. *Social Psychological and Personality Science*, *5*(1), 67–75. doi:10.1177/1948550613484499

Coleman, D. (2010). When Britain becomes "majority minority". Retrieved April 26, 2018, from https://www.prospectmagazine.co.uk/magazine/when-britain-becomes-majority-minority

Cooper, E. & Jahoda, M. (1947). The evasion of propaganda: How prejudiced people respond to anti-prejudice propaganda. *Journal of Psychology: Interdisciplinary and Applied*, *23*(1), 15–25. doi:10.1080/00223980.1947.9917316

Crisp, R. J. (2015). *The social brain: How diversity made the modern mind*. London, UK: Robinson.

Crisp, R. J., Birtel, M. D. & Meleady, R. (2011). Mental simulations of social thought and action. *Current Directions in Psychological Science*, *20*(4), 261–264. doi:10.1177/0963721411413762

Crisp, R. J. & Hewstone, M. (2007). Multiple social categorization. *Advances in Experimental Social Psychology*, *39*, 163–254. doi:10.1016/S0065-2601(06)39004-1

Crisp, R. J., Hewstone, M. & Rubin, M. (2001). Does multiple categorization reduce intergroup bias? *Personality and Social Psychology Bulletin*, *27*(1), 76–89. doi:10.1177/0146167201271007

Crisp, R. J. & Meleady, R. (2012). Adapting to a multicultural future. *Science*, *336*, 853–855. doi:10.1016/S0169-1317(03)00213-8

Crisp, R. J. & Turner, R. N. (2009). Can imagined interactions produce positive perceptions? Reducing prejudice through simulated social contact. *American Psychologist*, *64*(4), 231–240. doi:10.1037/a0014718

Crisp, R. J. & Turner, R. N. (2011). Cognitive adaptation to the experience of social and cultural diversity. *Psychological Bulletin*, *137*(2), 242–266. doi:10.1037/a0021840

Crisp, R. J. & Turner, R. N. (2012). The imagined contact hypothesis. *Advances in Experimental Social Psychology*, *46*, 125–182. doi:10.1016/B978-0-12-394281-4.00003-9

Crocker, J., Hannah, D. B. & Weber, R. (1983). Person memory and casual attributions. *Journal of Personality and Social Psychology*, *44*(1), 55–66. doi:http://dx.doi.org/10.1037/0022-3514.44.1.55

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. doi:10.1037/0033-2909.87.3.546

Crump, M. J., McDonnell, J. V. & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE*, *8*(3). doi:10.1371/journaltitle.pone.0057410

Darley, J. & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*(4), 377–383. doi:10.1037/h0025589

Dasgupta, N. & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology*, *40*(5), 642–658. doi:10.1016/j.jesp.2004.02.003

Deci, E. L. & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. doi:10.1207/S15327965PLI1104_01

Delhey, J. & Newton, K. (2005). Predicting cross-national levels of social trust: Global pattern or Nordic exceptionalism? *European Sociological Review*, *21*(4), 311–327. doi:10.1093/esr/jci022

Della Posta, D. J. (2013). Competitive threat, intergroup contact, or both? Immigration and the dynamics of Front National voting in France. *Social Forces*, *92*(1), 249–273. doi:10.1093/sf/sot046

Devine, P. G., Forscher, P. S., Austin, A. J. & Cox, W. T. L. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, *48*(6), 1267–1278. doi:10.1016/j.jesp.2012.06.003

Di Bella, L. & Crisp, R. J. (2016). Women's adaptation to stem domains promotes resilience and a lesser reliance on heuristic thinking. *Group Processes & Intergroup Relations*, *19*(2), 184–201. doi:10.1177/1368430215596074

Dickey, M. R. (2015). #ilooklikeanengineer aims to spread awareness about diversity in tech. Retrieved April 20, 2018, from https://techcrunch.com/2015/08/03/ilooklikeanengineer-aims-to-spread-awareness-about-gender-diversity-in-tech/

Dickter, C. & Gyurovski, I. (2012). The effects of expectancy violations on early attention to race in an impression-formation paradigm. *Social Neuroscience*, *7*(3), 240–251. doi:10.1080/17470919.2011.609906

Dollard, J., Miller, N. E., Doob, L. W., Mowrer, O. H. & Sears, R. R. (1939). *Frustration and aggression*. New Haven: Yale University Press. doi:10.1037/10022-000

Dovidio, J. F., Gaertner, S. L. & Kawakami, K. (2003). Intergroup contact: The past, present, and the future. *Group Processes & Intergroup Relations*, *6*(1), 5–21. doi:10.1177/1368430203006001009

Dovidio, J. F., Gaertner, S. L. & Loux, S. (2000). Subjective experiences and intergroup relations: The role of positive affect. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (Chap. 19, pp. 340–371). Philadelphia: Psychology Press.

Dovidio, J. F., Gaertner, S. L. & Saguy, T. (2009). Commonality and the complexity of "we": Social attitudes and social change. *Personality and Social Psychology Review*, *13*(1), 3–20. doi:10.1177/1088868308326751

Downey, H. K. & Slocum, J. W. (1975). Uncertainty: Measures, research, and sources of variation. *The Academy of Management Journal*, *18*(3), 562–578.

Dweck, C. S. (2017). From needs to goals and representations: Foundations for a unified theory of motivation, personality, and development. *Psychological Review*, *124*(6), 689–719. doi:10.1037/rev0000082

Eckel, C. C. & Grossman, P. J. (2001). Chivalry and solidarity in ultimatum games. *Economic Inquiry*, *39*(2), 171–188. doi:10.1093/ei/39.2.171

Eckstein, H. (1966). *Division and cohesion in democracy: A study of Norway*. Princeton, NJ: Princeton University Press.

Elliot, A. J. & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology*, *67*(3), 382–394. doi:10.1037/0022-3514.67.3.382

Ericsson, K. A., Krampe, T. T. & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. doi:http://dx.doi.org/10.1037/0033-295X.100.3.363

Esses, V. M., Dovidio, J. F., Jackson, L. M. & Armstrong, T. L. (2001). The immigration dilemma: The role of perceived group competition, ethnic prejudice, and national identity. *Journal of Social Issues*, *57*(3), 389–412. doi:10.1111/0022-4537.00220

Eurobarometer. (2012). *Europeans and their languages*. European Commission. Retrieved from http://ec.europa.eu/public%7B%5C_%7Dopinion/index%7B%5C_%7Den.htm

Evans, G. & Mellon, J. (2016). Working class votes and conservative losses: Solving the UKIP puzzle. *Parliamentary Affairs*, *69*(2), 464–479. doi:10.1093/pa/gsv005

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*(1), 255–278. doi:10.1146/annurev.psych.59.103006.093629

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. doi:10.3758/BF03193146

Fay, D. & Frese, M. (2000). Conservatives' approach to work: Less prepared for future work demands? *Journal of Applied Social Psychology*, *30*(1), 171–195. doi:https://doi.org/10.1111/j.1559-1816.2000.tb02310.x

Ferguson, C. J. & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, *7*(6), 555–561. doi:10.1177/1745691612459059

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press. doi:10.1037/10318-001

Fiedler, K. & Schwarz, N. (2016). Questionable research practices revisited. *Social Psychological and Personality Science*, *7*(1), 45–52. doi:10.1177/1948550615612150

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, *87*(6), 477–531. doi:http://dx.doi.org/10.1037/0033-295X.87.6.477

Fiske, S. T. & Neuberg, S. L. (1990). A continuum model of impression formation: From category-based to individuating processes as a function of information, motivation, and attention. *Advances in Experimental Social Psychology*, *23*, 1–74. doi:10.1016/S0065-2601(08)60317-2

Fiske, S. T. & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Fletcher, R. & Nielsen, R. K. (2017). *Using social media appears to diversify your news diet, not narrow it*. Reuters Institute for the Study of Journalism. Retrieved from http://www.niemanlab.org/2017/06/using-social-media-appears-to-diversify-your-news-diet-not-narrow-it/

Folkman, S., Lazarus, R. S., Dunkel-Schetter, C., DeLongis, A. & Gruen, R. J. (1986). Dynamics of a stressful encounter: Cognitive appraisal, coping, and

encounter outcomes. *Journal of Personality and Social Psychology*, *50*(5), 992–1003. doi:10.1037/0022-3514.50.5.992

Forscher, P. S., Mitamura, C., Dix, E. L., Cox, W. T. & Devine, P. G. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of Experimental Social Psychology*, *72*, 133–146. doi:10.1016/j.jesp.2017.04.009

Förster, J., Higgins, E. T. & Werth, L. (2004). How threat from stereotype disconfirmation triggers self-defense. *Social Cognition*, *22*(1), 54–74. doi:10.1521/soco.22.1.54.30982

Förster, J., Marguc, J. & Gillebaart, M. (2010). Novelty categorization theory. *Social and Personality Psychology Compass*, *4*(9), 736–755. doi:10.1111/j.1751-9004.2010.00289.x

Franzmann, S. (2017). A right-wing populist party founded by economists: The strange case of Germany's AfD. Retrieved December 13, 2017, from http://eprints.lse.ac.uk/69909/

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42. doi:10.1257/089533005775196732

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology. *American Psychologist*, *56*(3), 218–226. doi:10.1038/jid.2014.371

Fredrickson, B. L. & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition and Emotion*, *19*(3), 313–332. doi:10.1080/02699930441000238

Fredrickson, B. L. & Levenson, R. W. (1998). Positive emotions speed recovery from the cardiovascular sequelae of negative emotions. *Cognition and Emotion*, *12*(2), 191–220. doi:10.1080/026999398379718

Fritz, M. S. & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science*, *18*(3), 233–239. doi:10.1111/j.1467-9280.2007.01882.x

Fu, J. H.-y., Morris, M. W., Lee, S.-l., Chao, M., Chiu, C.-y. & Hong, Y.-y. (2007). Epistemic motives and cultural conformity: Need for closure, culture, and context as determinants of conflict judgments. *Journal of Personality and Social Psychology*, *92*(2), 191–207. doi:10.1037/0022-3514.92.2.191

Gaertner, S. L., Bachman, B. A., Dovidio, J. D. & Banker, B. S. (2001). Corporate mergers and stepfamily marriages: Identity, harmony, and commitment. In M. A. Hogg & D. J. Terry (Eds.), *Social identity processes in organizational contexts* (Chap. 17, pp. 365–288). Oxford, England: Psychology Press.

Gaertner, S. L. & Bickman, L. (1971). Effects of race on the elicitation of helping behavior: The wrong number technique. *Journal of Personality and Social Psychology*, *20*(2), 218–222. doi:10.1037/h0031681

Gaertner, S. L., Mann, J. A., Dovidio, J. F., Murrell, A. J. & Pomare, M. (1990). How does cooperation reduce intergroup bias? *Journal of Personality and Social Psychology*, *59*(4), 692–704. doi:10.1037/0022-3514.59.4.692

Gaertner, S. L., Rust, M. C., Dovidio, J. F., Bachman, B. A. & Anastasio, P. A. (1994). The contact hypothesis: The role of a common ingroup identity in

reducing intergroup bias. *Small Group Research*, *25*(2), 224–249. doi:10.1177/07399863870092005

Galinsky, A. D., Todd, A. R., Homan, A. C., Phillips, K. W., Apfelbaum, E. P., Sasaki, S. J., ... Maddux, W. W. (2015). Maximizing the gains and minimizing the pains of diversity. *Perspectives on Psychological Science*, *10*(6), 742–748. doi:10.1177/1745691615598513

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., ... Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, *332*(6033), 1100–1104. doi:10.1126/science.1197754

Gelman, A. & Loken, E. (2014). The statistical crisis in science. *American Scientist*, *102*(6), 460. doi:10.1511/2014.111.460

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, *3*(1), 20–29. doi:https://doi.org/10.1111/j.1745-6916.2008.00058.x

Gilbert, D. T., King, G., Pettigrew, S. & Wilson, T. D. (2016). Comment on "Estimating the reproducibility of psychological science". *Science*, *351*(6277), 1037–1037. doi:10.1126/science.aad7243

Gilbert, E. A. & Corker, K. S. (2017). Research transparency: 5 questions about open science answered. *The Conversation*. Retrieved from https://theconversation.com/research-transparency-5-questions-about-open-science-answered-76851

Giles, M. W. & Evans, A. (1986). The power approach to intergroup hostility. *Journal of Conflict Resolution*, *30*(3), 469–486. doi:https://doi.org/10.1177/0022002786030003004

Gino, F. & Ariely, D. (2012). The dark side of creativity: Original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, *102*(3), 445–459. doi:10.1037/a0026406

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A. & Soutter, C. L. (2000). Measuring trust. *Quarterly Journal of Economics*, *115*(3), 811–846. doi:10.1162/003355300554926

Gocłowska, M. A., Baas, M., Crisp, R. J. & De Dreu, C. K. W. (2014). Whether social schema violations help or hurt creativity depends on need for structure. *Personality and Social Psychology Bulletin*, *40*(8), 959–971. doi:10.1177/0146167214533132

Gocłowska, M. A., Baas, M., Elliot, A. J. & De Dreu, C. K. (2017a). Why schema-violations are sometimes preferable to schema-consistencies: The role of interest and openness to experience. *Journal of Research in Personality*, *66*, 54–69. doi:10.1016/j.jrp.2016.12.005

Gocłowska, M. A. & Crisp, R. J. (2013). On counter-stereotypes and creative cognition: When interventions for reducing prejudice can boost divergent thinking. *Thinking Skills and Creativity*, *8*(1), 72–79. doi:10.1016/j.tsc.2012.07.001

Gocłowska, M. A., Crisp, R. J. & Labuschagne, K. (2012). Can counter-stereotypes boost flexible thinking? *Group Processes & Intergroup Relations*, *16*(2), 217–231. doi:10.1177/1368430212445076

Gocłowska, M. A., Damian, R. I. & Mor, S. (2017b). The diversifying experience model: Taking a broader conceptual view of the multiculturalism–

creativity link. *Journal of Cross-Cultural Psychology, in press*. doi:10.1177/0022022116650258

Graham, J., Haidt, J. & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029–1046. doi:10.1037/a0015141

Graham, J., Nosek, B. A. & Haidt, J. (2012). The moral stereotypes of liberals and conservatives: Exaggeration of differences across the political spectrum. *PLoS ONE, 7*(12). doi:10.1371/journaltitle.pone.0050092

Greenberg, J., Porteus, J., Simon, L., Pyszczynski, T. & Solomon, S. (1995). Evidence of a terror management function of cultural icons: The effects of mortality salience on the inappropriate use of cherished cultural symbols. *Personality and Social Psychology Bulletin, 21*(11), 1221–1228. doi:10.1177/01461672952111010

Greenberg, J., Solomon, S., Veeder, M., Pyszczynski, T., Rosenblatt, A., Kirkland, S. & Lyon, D. (1990). Evidence for terror management theory II: The effects of mortality salience on reactions to those who threaten or bolster the cultural worldview. *Journal of Personality and Social Psychology, 58*(2), 308–318. doi:10.1037/0022-3514.58.2.308

Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464–80. doi:10.1037/0022-3514.74.6.1464

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. doi:10.1037/0033-295X.109.1.3

Griskevicius, V., Shiota, M. N. & Neufeld, S. L. (2010). Influence of different positive emotions on persuasion processing: A functional evolutionary approach. *Emotion*, *10*(2), 190–206. doi:10.1037/a0018421

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, *2*(5), 271–299. doi:10.1037/1089-2680.2.3.271

Hall, N. R. & Crisp, R. J. (2005). Considering multiple criteria for social categorization can reduce intergroup bias. *Personality and Social Psychology Bulletin*, *31*(10), 1435–1444. doi:10.1177/0146167205276084

Halperin, E., Crisp, R. J., Husnu, S., Trzesniewski, K. H., Dweck, C. S. & Gross, J. J. (2012). Promoting intergroup contact by changing beliefs: Group malleability, intergroup anxiety, and contact motivation. *Emotion*, *12*(6), 1192–1195. doi:10.1037/a0028620

Hameiri, B., Porat, R., Bar-Tal, D., Bieler, A. & Halperin, E. (2014). Paradoxical thinking as a new avenue of intervention to promote peace. *Proceedings of the National Academy of Sciences*, *111*(30), 10996–11001. doi:10.1073/pnas.1407055111

Harmon-Jones, E. (2000). Cognitive dissonance and experienced negative affect: Evidence that dissonance increases experienced negative affect even in the absence of aversive consequences. *Personality and Social Psychology Bulletin*, *26*(12), 1490–1501. doi:10.1177/01461672002612004

Harmon-Jones, E., Amodio, D. M. & Harmon-Jones, C. (2009). Action-based model of dissonance: A review, integration, and expansion of conceptions of cognitive conflict. *Advances in Experimental Social Psychology*, *41*, 119–166. doi:10.1016/S0065-2601(08)00403-6

Hastie, R. & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, *37*(1), 25–38. doi:10.1037/0022-3514.37.1.25

Hastie, R., Schroeder, C. & Weber, R. (1990). Creating complex social conjunction categories from simple categories. *Bulletin of the Psychonomic Society*, *28*(3), 242–247. doi:10.3758/BF03334016

Hayward, L. E., Tropp, L. R., Hornsey, M. J. & Barlow, F. K. (2017). Toward a comprehensive understanding of intergroup contact: Descriptions and mediators of positive and negative contact among majority and minority groups. *Personality and Social Psychology Bulletin*, *43*(3), 347–364. doi:10.1177/0146167216685291

Heider, F. (1958). *The psychology of interpersonal relations*. Hoboken, NJ, US: John Wiley & Sons Inc. doi:http://dx.doi.org/10.1037/10628-000

Heider, F. & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*, 243–249. doi:10.2307/1416950

Heine, S. J., Proulx, T. & Vohs, K. D. (2006). The Meaning Maintenance Model: On the coherence of social motivations. *Personality and Social Psychology Review*, *10*(2), 88–110. doi:10.1207/s15327957pspr1002_1

Herring, C. (2009). Does diversity pay? Race, gender, and the business case for diversity. *American Sociological Review*, *74*(2), 208–224. doi:10.1177/000312240907400203

Hewstone, M. (1994). Revision and change of stereotypic beliefs: In search of the elusive subtyping model. *European Review of Social Psychology*, *5*(1), 69–109. doi:10.1080/14792779543000020

Hewstone, M. & Brown, R. (1986). Contact is not enough: An intergroup perspective on the 'contact hypothesis.' In M. Hewstone & R. Brown (Eds.), *Social psychology and society. contact and conflict in intergroup encounters* (pp. 1–44). Cambridge, MA: Basil Blackwell.

Hewstone, M., Rubin, M. & Willis, H. (2001). Intergroup bias. *Annual Review of Psychology*, *53*(1), 575–604. doi:https://doi.org/10.1146/annurev.psych.53.100901.135109

Hewstone, M. & Swart, H. (2011). Fifty-odd years of inter-group contact: From hypothesis to integrated theory. *British Journal of Social Psychology*, *50*, 374–386. doi:10.1111/j.2044-8309.2011.02047.x

Hoever, I. J., van Knippenberg, D., van Ginkel, W. P. & Barkema, H. G. (2012). Fostering team creativity: Perspective taking as key to unlocking diversity's potential. *Journal of Applied Psychology*, *97*(5), 982–996. doi:10.1037/a0029159

Hoffarth, M. R. & Hodson, G. (2016). Who needs imagined contact? Replication attempts examining previous contact as a potential moderator. *Social Psychology*, *47*(2), 118–124. doi:10.1027/1864-9335/a000258

Hoffman, L. R. (1959). Homogeneity of member personality and its effect on group problem-solving. *Journal of Abnormal Psychology*, *58*(1), 27–32. doi:10.1037/h0043499

Hofstede, G. (1983). National cultures in four dimensions: A research-based theory of cultural differences among nations. *International Studies of Management & Organization*, *13*(1-2), 46–74. doi:10.1080/00208825.1983.11656358

Hogg, M. A. (2000). Subjective uncertainty reduction through self-categorization: A motivational theory of social identity processes. *European Review of Social Psychology*, *11*(1), 223–255. doi:10.1080/14792772043000040

Hogg, M. A. (2007). Uncertainty-identity theory. *Advances in Experimental Social Psychology*, *39*, 69–126. doi:10.1016/S0065-2601(06)39002-8

Hohwy, J. (2014). *The predictive mind*. Oxford, England: Oxford University Press.

Hornsey, M. J. & Hogg, M. A. (2000). Assimilation and diversity: An integrative model of subgroup relations. *Personality and Social Psychology Review*, *4*(2), 143–156. doi:https://doi.org/10.1207/S15327957PSPR0402_03

Horstmann, G. (2015). The surprise-attention link: A review. *Annals of the New York Academy of Sciences*, *1339*(1), 106–115. doi:10.1111/nyas.12679

Houlette, M. A., Gaertner, S. L., Johnson, K. M., Banker, B. S., Riek, B. M. & Dovidio, J. F. (2004). Developing a more inclusive social identity: An elementary school intervention. *Journal of Social Issues*, *60*(1), 35–55. doi:10.1111/j.0022-4537.2004.00098.x

Hovland, C. I. & Sears, R. R. (1940). Minor studies of aggression: VI. Correlation of lynchings with economic indices. *The Journal of Psychology*, *9*(2), 301–310. doi:10.1080/00223980.1940.9917696

Husnu, S. & Crisp, R. J. (2010). Elaboration enhances the imagined contact effect. *Journal of Experimental Social Psychology*, *46*(6), 943–950. doi:10.1016/j.jesp.2010.05.014

Hutter, R. R. C. & Crisp, R. J. (2005). The composition of category conjunctions. *Personality and Social Psychology Bulletin*, *31*(5), 647–657. doi:10.1177/0146167204271575

Hutter, R. R. C. & Crisp, R. J. (2006). Implications of cognitive busyness for the perception of category conjunctions. *The Journal of Social Psychology*, *146*(2), 253–256. doi:10.3200/SOCP.146.2.253-256

Hutter, R. R. C., Crisp, R. J., Humphreys, G. W., Waters, G. M. & Moffitt, G. (2009). The dynamics of category conjunctions. *Group Processes & Intergroup Relations*, *12*(5), 673–686. doi:10.1177/1368430209337471

Hutter, R. R. C., Wood, C. & Dodd, G. F. (2012). Resolving conflicting social categories: The role of age-related executive ability. *British Journal of Psychology*, *103*(1), 28–43. doi:10.1111/j.2044-8295.2011.02037.x

Ibrahim, J. G. & Molenberghs, G. (2009). Missing data methods in longitudinal studies: A review. *Test*, *18*(1), 1–43. doi:10.1007/s11749-009-0138-x

Inbar, Y., Pizarro, D. A. & Bloom, P. (2009). Conservatives are more easily disgusted than liberals. *Cognition and Emotion*, *23*(4), 714–725. doi:10.1080/02699930802110007

Inzlicht, M. & Schmeichel, B. J. (2013). Beyond simple utility in predicting self-control fatigue: A proximate alternative to the opportunity cost model. *Behavioral and Brain Sciences*, *36*(06), 695–696. doi:10.1017/S0140525X13001076

Isen, A. M., Daubman, K. A. & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, *52*(6), 1122–1131. doi:10.1037/0022-3514.52.6.1122

Jackson, S. E., Brett, J. F., Sessa, V. I., Cooper, D. M., Julin, J. A. & Peyronnin, K. (1991). Some differences make a difference: Individual dissimilarity and group heterogeneity as correlates of recruitment, promotions, and turnover. *Journal of Applied Psychology*, *76*(5), 675–689. doi:10.1037/0021-9010.76.5.675

Janis, I. L. (1972). *Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes.* Oxford, England: Houghton Mifflin.

John, L. K., Loewenstein, G. & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, *23*(5), 524–532. doi:10.1177/0956797611430953

Johns, N., Green, A. & Powell, M. (2012). Diversity in the British NHS: The business versus the "moral" case. *Equality, Diversity and Inclusion: An International Journal*, *31*(8), 768–783. doi:10.1108/02610151211277626

Johnston, L. & Hewstone, M. (1992). Cognitive models of stereotype change: 3. Subtyping and the perceived typicality of disconfirming group members. *Journal of Experimental Social Psychology*, *28*(4), 360–386. doi:10.1016/0022-1031(92)90051-K

Jonas, E., McGregor, I., Klackl, J., Agroskin, D., Fritsche, I., Holbrook, C., ... Quirin, M. (2014). Threat and defense: From anxiety to approach. *Advances in Experimental Social Psychology*, *49*, 219–286. doi:10.1016/B978-0-12-800052-6.00004-4

Jones, E. E. & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. *Advances in Experimental Social Psychology*, *2*, 219–266. doi:10.1016/S0065-2601(08)60107-0

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. doi:10.1111/j.1467-9221.2004.00402.x

Jost, J. T., Federico, C. M. & Napier, J. L. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual Review of Psychology*, *60*(1), 307–337. doi:10.1146/annurev.psych.60.110707.163600

Jost, J. T., Glaser, J., Kruglanski, A. W. & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339–375. doi:10.1037/0033-2909.129.3.339

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

Jussim, L., Coleman, L. M. & Lerch, L. (1987). The nature of stereotypes: A comparison and integration of three theories. *Journal of Personality and Social Psychology*, *52*(3), 536–546. doi:10.1037/0022-3514.52.3.536

Kahneman, D. (2012). *Thinking, fast and slow*. London, UK: Penguin Books.

Kahneman, D., Knetsch, J. L. & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The Journal of Economic Perspectives*, *5*(1), 193–206. doi:10.2307/1942711

Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*(4), 237–251. doi:http://dx.doi.org/10.1037/h0034747

Kassin, S. M. (1979). Consensus information, prediction, and causal attribution: A review of the literature and issues. *Journal of Personality and Social Psychology*, *37*(11), 1966–1981. doi:10.1037/0022-3514.37.11.1966

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S. & Russin, a. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, *78*(5), 871–888. doi:10.1037/0022-3514.78.5.871

Keeley, B. (2007). *Human capital: How what you know shapes your life*. Paris, France: Organisation for Economic Co-operation and Development.

Keller, R. T. (2001). Cross-functional groups in research and new project product development: Diversity, communications, job stress, and outcomes. *Academy of Management Journal*, *44*(3), 547–555. doi:10.2307/3069369

Kelley, H. (1967). Attribution theory in social psychology. *Nebraska Symposium on Motivation*, *15*, 192–238.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*(2), 107–128. doi:10.1037/h0034225

Kelley, H. H. & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, *31*(1), 457–501. doi:10.1146/annurev.ps.31.020180.002325

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, *2*(3), 196–217. doi:https://doi.org/10.1207/s15327957pspr0203_4

Kerschreiter, R., Schulz-Hardt, S., Mojzisch, A. & Frey, D. (2008). Biased information search in homogeneous groups: Confidence as a moderator for the effect of anticipated task requirements. *Personality and Social Psychology Bulletin*, *34*(5), 679–691. doi:10.1177/0146167207313934

Killian, L. M., Festinger, L., Riecken, H. W. & Schachter, S. (1957). When prophecy fails. *American Sociological Review*, *22*(2), 236. doi:10.2307/2088869

Kleiman, T. & Hassin, R. R. (2013). When conflicts are good: Nonconscious goal conflicts reduce confirmatory thinking. *Journal of Personality and Social Psychology*, *105*(3), 374–387. doi:10.1037/a0033608

Kobler, B. (2017). *International migration report 2017: Highlights*. United Nations. New York, NY.

Koch, A., Imhoff, R., Dotsch, R., Unkelbach, C. & Alves, H. (2016). The ABC of stereotypes about groups: Agency/socioeconomic success, conservative–progressive beliefs, and communion. *Journal of Personality and Social Psychology*, *110*(5), 675–709. doi:10.1037/pspa0000046

Kruglanski, A. W., Pierro, A., Mannetti, L. & De Grada, E. (2006). Groups as epistemic providers: Need for closure and the unfolding of group-centrism. *Psychological Review*, *113*(1), 84–100. doi:10.1037/0033-295X.113.1.84

Kruglanski, A. W. & Webster, D. M. (1991). Group members' reactions to opinion deviates and conformists at varying degrees of proximity to decision deadline and of environmental noise. *Journal of Personality and Social Psychology*, *61*(2), 212–225. doi:10.1037/0022-3514.61.2.212

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. doi:10.1037/0033-2909.108.3.480

Kunda, Z., Miller, D. T. & Claire, T. (1990). Combining social concepts: The role of causal reasoning. *Cognitive Science*, *14*(4), 551–577. doi:10.1016/0364-0213(90)90009-L

Kunda, Z. & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology*, *68*(4), 565–579. doi:10.1037/0022-3514.68.4.565

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., Joy-Gaba, J. A., ... Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, *143*(4), 1765–1785. doi:10.1037/a0036260

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., ... Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*(8), 1001–1016. doi:10.1037/xge0000179

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *The American Psychologist*, *46*(8), 819–834. doi:10.1037/0003-066X.46.8.819

Lazarus, R. S. & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality*, *1*(3), 141–169. doi:10.1002/per.2410010304

Leary, M. R., Diebels, K. J., Davisson, E. K., Jongman-Sereno, K. P., Isherwood, J. C., Raimi, K. T., ... Hoyle, R. H. (2017). Cognitive and interpersonal features of intellectual humility. *Personality and Social Psychology Bulletin*, *43*(6), 793–813. doi:10.1177/0146167217697695

Legault, L., Gutsell, J. N. & Inzlicht, M. (2011). Ironic effects of antiprejudice messages. *Psychological Science*, *22*(12), 1472–1477. doi:10.1177/0956797611427918

Leone, C., Wallace, H. M. & Modglin, K. (1999). The need for closure and the need for structure: Interrelationships, correlates, and outcomes. *Journal of Psychology: Interdisciplinary and Applied*, *133*(5), 553–562. doi:10.1080/00223989909599762

Leung, A. K. & Chiu, C. Y. (2008). Interactive effects of multicultural experiences and openness to experience on creative potential. *Creativity Research Journal*, *20*(4), 376–382. doi:10.1080/10400410802391371

LeVine, R. A. & Campbell, D. T. (1972). *Ethnocentrism: Theories of conflict, ethnic attitudes, and group behavior*. Oxford, England: John Wiley & Sons.

Li, J. (2005). Mind or virtue: Western and Chinese beliefs about learning. *Current Directions in Psychological Science*, *14*(4), 190–194. doi:https://doi.org/10.1111/j.0963-7214.2005.00362.x

Linville, P. W. (1987). Self-complexity as a cognitive buffer against stress-related illness and depression. *Journal of Personality and Social Psychology*, *52*(4), 663–676. doi:10.1037/0022-3514.52.4.663

Litman, J. A. (2008). Interest and deprivation factors of epistemic curiosity. *Personality and Individual Differences*, *44*(7), 1585–1595. doi:10.1016/j.paid.2008.01.014

Litman, J. A., Crowson, H. M. & Kolinski, K. (2010). Validity of the interest- and deprivation-type epistemic curiosity distinction in non-students. *Personality and Individual Differences*, *49*(5), 531–536. doi:10.1016/j.paid.2010.05.021

Litman, J. A. & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diversive and specific components. *Journal of Personality Assessment*, *80*(1), 75–86. doi:10.1207/S15327752JPA8001_16

Lundvall, B.-Å. (2002). *Innovation, growth and social cohesion: The Danish model*. Cheltenham, UK: Edward Elgar Publishing.

Macrae, C. N., Bodenhausen, G. V., Schloerscheidt, A. M. & Milne, A. B. (1999). Tales of the unexpected: Executive function and person perception. *Journal of Personality and Social Psychology*, *76*(2), 200–213. doi:10.1037/0022-3514.76.2.200

Maddux, W. W., Adam, H. & Galinsky, A. D. (2010). When in Rome ... learn why the Romans do what they do: How multicultural learning experiences

facilitate creativity. *Personality and Social Psychology Bulletin*, *36*(6), 731–741. doi:10.1177/0146167210367786

Maddux, W. W., Bivolaru, E., Hafenbrack, A. C., Tadmor, C. T. & Galinsky, A. D. (2014). Expanding opportunities by opening your mind: Multicultural engagement predicts job market success through longitudinal increases in integrative complexity. *Social Psychological and Personality Science*, *5*(5), 608–615. doi:10.1177/1948550613515005

Major, B., Blodorn, A. & Major Blascovich, G. (2016). The threat of increasing diversity: Why many White Americans support Trump in the 2016 presidential election. *Group Processes & Intergroup Relations*, Advance online publication. doi:10.1177/1368430216677304

Markus, H. R. & Kitayama, S. (1994). A collective fear of the collective: Implications for selves and theories of selves. *Personality and Social Psychology Bulletin*, *20*(5), 568–579. doi:10.1177/0146167294205013

Martens, P., Dreher, A. & Gaston, N. (2010). Globalisation, the global village and the civil society. *Futures*, *42*(6), 574–582. doi:10.1016/J.FUTURES.2010.01.008

Marx, D. M., Ko, S. J. & Friedman, R. A. (2009). The "Obama effect": How a salient role model reduces race-based performance differences. *Journal of Experimental Social Psychology*, *45*(4), 953–956. doi:10.1016/j.jesp.2009.03.012

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370–396. doi:http://dx.doi.org/10.1037/h0054346

Mayer, J. D. & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology*, *55*(1), 102–111. doi:10.1037/0022-3514.55.1.102

Mayseless, O. & Kruglanski, A. W. (1987). What makes you so sure? Effects of epistemic motivations on judgmental confidence. *Organizational Behavior and Human Decision Processes*, *39*(2), 162–183. doi:10.1016/0749-5978(87)90036-7

Mcarthur, L. A. (1972). The how and what of why: Some determinants and consequences of causal attribution. *Journal ol Personality and Social Psychology*, *22*(2), 171–193. doi:http://dx.doi.org/10.1037/h0032602

McFalls, E. L. & Cobb-Roberts, D. (2001). Reducing resistance to diversity through cognitive dissonance instruction. *Journal of Teacher Education*, *52*(2), 164–172. doi:10.1177/0022487101052002007

McGregor, I., Zanna, M. P., Holmes, J. G. & Spencer, S. J. (2001). Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology*, *80*(3), 472–488. doi:10.1037/0022-3514.80.3.472

Meleady, R. & Crisp, R. J. (2014). Systematic revisions to inherent notions may shape improvements in cognitive infrastructure. *Behavioral and Brain Sciences*, *37*(5), 495–496. doi:10.1017/S0140525X13003804

Meleady, R., Seger, C. R. & Vermue, M. (2017). Examining the role of positive and negative intergroup contact and anti-immigrant prejudice in Brexit. *British Journal of Social Psychology*, *56*(4), 1–10. doi:10.1111/bjso.12203

Mendes, W. B., Blascovich, J., Hunter, S. B., Lickel, B. & Jost, J. T. (2007). Threatened by the unexpected: Physiological responses during social interactions with expectancy-violating partners. *Journal of Personality and Social Psychology*, *92*(4), 698–716. doi:10.1037/0022-3514.92.4.698

Mendoza-Denton, R. & España, C. (2010). Diversity science: What is it? *Psychological Inquiry*, *21*(2), 140–145. doi:10.1080/10478401003800945

Miles, E. & Crisp, R. J. (2014). A meta-analytic test of the imagined contact hypothesis. *Group Processes & Intergroup Relations*, *17*(1), 3–26. doi:10.1177/1368430213510573

Milgram, S. (1970). The experience of living in cities. *Science*, *167*(3924), 1461–1468. doi:http://www.jstor.org/stable/1728966

Mirisola, A. A. & Seta, L. (2016). pequod: Moderated regression package. Retrieved from http://cran.r-project.org/package=pequod

Mogg, K., McNamara, J., Rawlinson, H., Seiffer, A., Bradley, B. P., Powys, M., ... Bradley, B. P. (2000). Selective attention to threat: A test of two cognitive models of anxiety. *Cognition & Emotion*, *14*(3), 375–399. doi:10.1080/026999300378888

Molinsky, A. (2007). Cross-cultural code-switching: The psychological challenges of adapting behavior in foreign cultural interactions. *Academy of Management Review*, *32*(2), 622–640. doi:http://dx.doi.org/10.2307/20159318

Morris, J. A. & Feldman, D. C. (1996). The dimensions, antecedents, and consequences of emotional labor. *Academy of Management Review*, *21*(4), 986–1010. doi:10.5465/AMR.1996.9704071861

Moss-Racusin, C. A., van der Toorn, J., Dovidio, J. F., Brescoll, V. L., Graham, M. J. & Handelsman, J. (2014). Scientific diversity interventions. *Science*, *343*(6171), 615–616. doi:10.1126/science.1245936

Mu, Y., Kitayama, S., Han, S. & Gelfand, M. J. (2015). How culture gets embrained: Cultural differences in event-related potentials of social norm violations. *Proceedings of the National Academy of Sciences*, *112*(50), 15348–15353. doi:10.1073/pnas.1509839112

Mullainathan, S. & Shafir, E. (2013). *Scarcity: Why having too little means so much*. New York, NY: Holt.

Munafò, M. R. & Smith, G. D. (2018). Robust research needs many lines of evidence. *Nature*, *553*(7689), 399–401. doi:10.1038/d41586-018-01023-3

Mussel, P. (2010). Epistemic curiosity and related constructs: Lacking evidence of discriminant validity. *Personality and Individual Differences*, *49*(5), 506–510. doi:10.1016/j.paid.2010.05.014

Nair, K. U. & Ramnarayan, S. (2000). Individual Differences in need for cognition and complex problem solving. *Journal of Research in Personality*, *34*(3), 305–328. doi:10.1006/JRPE.1999.2274

Nelson, L. D., Simmons, J. & Simonsohn, U. (2018). Psychology's renaissance. *Annual Review of Psychology*, *69*(17), 1–24. doi:https://doi.org/10.1146/annurev-psych-122216-011836

Neto, F., Pinto, M. D. C. & Mullet, E. (2015). Can music reduce anti-dark-skin prejudice? A test of a cross-cultural musical education programme. *Psychology of Music*, *44*(3), 388–398. doi:10.1177/0305735614568882

Neuberg, S. L. & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality and Social Psychology*, *65*(1), 113–131. doi:10.1037/0022-3514.65.1.113

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. doi:10.1037/1089-2680.2.2.175

Nier, J. a., Gaertner, S. L., Dovidio, J. F., Banker, B. S., Ward, C. M. & Rust, M. C. (2001). Changing interracial evaluations and behavior: The effects of a common group identity. *Group Processes & Intergroup Relations*, *4*(4), 299–316. doi:https://doi.org/10.1177/1368430201004004001

Noon, M. (2007). The fatal flaws of diversity and the business case for ethnic minorities. *Work, Employment and Society*, *21*(4), 773–784. doi:10.1177/0950017007082886

Noordewier, M. K. & Breugelmans, S. M. (2013). On the valence of surprise. *Cognition and Emotion*, *27*(7), 1326–1334. doi:10.1080/02699931.2013.777660

Noordewier, M. K., Topolinski, S. & Dijk, E. V. (2016). The temporal dynamics of surprise. *Social and Personality Psychology Compass*, *103*(10), 136–149. doi:10.1111/spc3.12242

Nordgren, L. F., van Harreveld, F. & van der Pligt, J. (2006). Ambivalence, discomfort, and motivated information processing. *Journal of Experimental Social Psychology*, *42*(2), 252–258. doi:10.1016/j.jesp.2005.04.004

Norenzayan, A. & Hansen, I. G. (2006). Belief in supernatural agents in the face of death. *Personality and Social Psychology Bulletin*, *32*(2), 174–187. doi:10.1177/0146167205280251

Norman, D. A. & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self regulation: Advances in research* (Vol. 4). New York, NY: Plenum Press. doi:http://sci-hub.tw/10.1007/978-1-4757-0629-1_1

Nosek, B. A., Ebersole, C. R., DeHaven, A. C. & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. doi:10.1073/pnas.1708274114

Notebaert, L., Crombez, G., Van Damme, S., De Houwer, J. & Theeuwes, J. (2011). Signals of threat do not capture, but prioritize, attention: A conditioning approach. *Emotion*, *11*(1), 81–89. doi:10.1037/a0021286

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), 943–952. doi:10.1126/science.aac4716

Oppenheimer, D. M., Meyvis, T. & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872. doi:10.1016/j.jesp.2009.03.009

Ortman, J. M. & Guarneri, C. E. (2009). *United States population projections: 2000 to 2050*. U.S. Census Bureau. Retrieved from https://census.gov/content/dam/Census/library/working-papers/2009/demo/us-pop-proj-2000-2050/analytical-document09.pdf

Packer, G. (2014). The quiet German: The astonishing rise of Angela Merkel, the most powerful woman in the world. Retrieved April 22, 2018, from https://www.newyorker.com/magazine/2014/12/01/quiet-german

Palan, S. & Schitter, C. (2018). Prolific.ac – a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. doi:10.1016/j.jbef.2017.12.004

Paluck, E. L. & Green, D. P. (2009). Prejudice reduction: What works? a review and assessment of research and practice. *Annual Review of Psychology*, *60*(1), 339–367. doi:10.1146/annurev.psych.60.110707.163607

Pan, C., Pfeil, B. S. & Videsott, P. (2016). *Handbuch der europäischen Volksgruppen Band 1: Die Volksgruppen in Europa* (2nd ed.). Berliner Wissenschafts-Verlag. doi:http://hdl.handle.net/10863/2738

Pariser, E. (2015). Did Facebook's big study kill my filter bubble thesis? *Wired*. Retrieved from https://www.wired.com/2015/05/did-facebooks-big-study-kill-my-filter-bubble-thesis/

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. doi:10.1207/s15327957pspr0902_2

Park, C. L. & Folkman, S. (1997). Meaning in the context of stress and coping. *Review of General Psychology*, *1*(2), 115–144. doi:10.1037/1089-2680.1.2.115

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S. & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for aca-

demic underachievement. *Psychological Science*, *26*(6), 784–793. doi:10.1177/0956797615571017

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation into the physiological activity of the cortex*. New York, NY: Dover.

Peer, E., Brandimarte, L., Samat, S. & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. doi:10.1016/j.jesp.2017.01.006

Pelled, L. H., Eisenhardt, K. M. & Xin, K. R. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly*, *44*(1), 1. doi:10.2307/2667029

Peltier, J. W. & Schibrowsky, J. A. (1994). Need for cognition, advertisement viewing time and memory for advertising stimuli. *Advances in Consumer Research*, *21*, 244–250. doi:http://acrwebsite.org/volumes/7595/volumes/v21/NA-21

Perlovsky, L. & Ilin, R. (2012). Brain. Conscious and unconscious mechanisms of cognition, emotions, and language. *Brain Sciences*, *2*(4), 790–834. doi:10.3390/brainsci2040790

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, *49*(1), 65–85. doi:10.1146/annurev.psych.49.1.65

Pettigrew, T. F. (2008). Future directions for intergroup contact theory and research. *International Journal of Intercultural Relations*, *32*(3), 187–199. doi:10.1016/j.ijintrel.2007.12.002

Pettigrew, T. F. & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. doi:10.1037/0022-3514.90.5.751

Pew Research Center. (2016). 5 facts about migration and the United Kingdom. Retrieved April 27, 2018, from http://www.pewresearch.org/fact-tank/2016/06/21/5-facts-about-migration-and-the-united-kingdom/

Piaget, J. (1953). *The origins of intelligence in children*. New York, NY: International University Press.

Piaget, J. (1964). Part I: Cognitive development in children. *Journal of Research in Science Teaching*, *2*(3), 176–186. doi:10.1002/tea.3660020306

Plaks, J. E., Grant, H. & Dweck, C. S. (2005). Violations of implicit theories and the sense of prediction and control: Implications for motivated person perception. *Journal of Personality and Social Psychology*, *88*(2), 245–262. doi:10.1037/0022-3514.88.2.245

Plaks, J. E., Stroessner, S. J., Dweck, C. S. & Sherman, J. W. (2001). Person theories and attention allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of Personality & Social Psychology*, *80*(6), 876–893. doi:10.1037/0022-3514.80.6.876

Plaut, V. C. (2010a). Diversity science: Who needs it? *Psychological Inquiry*, *21*(2), 168–174. doi:10.1080/1047840X.2010.492753

Plaut, V. C. (2010b). Diversity science: Why and how difference makes a difference. *Psychological Inquiry*, *21*(2), 77–99. doi:10.1080/10478401003676501

Prati, F., Crisp, R. J., Pratto, F. & Rubini, M. (2016). Encouraging majority support for immigrant access to health services: Multiple categorization and social identity complexity as antecedents of health equality. *Group Processes & Intergroup Relations*, *19*(4), 426–438. doi:10.1177/1368430216629814

Prati, F., Crisp, R. J. & Rubini, M. (2015a). Counter-stereotypes reduce emotional intergroup bias by eliciting surprise in the face of unexpected category combinations. *Journal of Experimental Social Psychology*, *61*, 31–43. doi:10.1016/j.jesp.2015.06.004

Prati, F., Vasiljevic, M., Crisp, R. J. & Rubini, M. (2015b). Some extended psychological benefits of challenging social stereotypes: Decreased dehumanization and a reduced reliance on heuristic thinking. *Group Processes & Intergroup Relations*, *18*(6), 801–816. doi:10.1177/1368430214567762

Proulx, T. & Heine, S. J. (2006). Death and black diamonds: Meaning, mortality, and the meaning maintenance model. *Psychological Inquiry*, *17*(4), 309–318. doi:10.1080/10478400701366985

Proulx, T. & Inzlicht, M. (2012). The five "A"s of meaning maintenance: Finding meaning in the theories of sense-making. *Psychological Inquiry*, *23*(4), 317–335. doi:10.1080/1047840X.2012.702372

Proulx, T., Inzlicht, M. & Harmon-Jones, E. (2012). Understanding all inconsistency compensation as a palliative response to violated expectations. *Trends in Cognitive Sciences*, *16*(5), 285–291. doi:10.1016/j.tics.2012.04.002

Psychology. (2018). English Oxford Dictionaries. Retrieved from https://en.oxforddictionaries.com/definition/psychology

Pullman, P. (2000). *The Amber Spyglass*. London, UK: David Fickling.

Putnam, R. D. (2007). E pluribus unum: Diversity and community in the twenty-first century the 2006 Johan Skytte Prize Lecture. *Scandinavian Political Studies*, *30*(2), 137–174. doi:10.1111/j.1467-9477.2007.00176.x

Quattrone, G. A. & Jones, E. E. (1980). The perception of variability within ingroups and out-groups: Implications for the law of small numbers. *Journal of Personality and Social Psychology*, *38*(1), 141–152. doi:10.1037/0022-3514.38.1.141

Quillian, L. (1995). Prejudice as a response to perceived group threat: Population composition and anti-immigrant and racial prejudice in Europe. *American Sociological Review*, *60*(4), 586–611. doi:10.2307/2096296

Randles, D., Inzlicht, M., Proulx, T., Tullett, A. M. & Heine, S. J. (2015). Is dissonance reduction a special case of fluid compensation? Evidence that dissonant cognitions cause compensatory affirmation and abstraction. *Journal of Personality and Social Psychology*, *108*(5), 697–710. doi:10.1037/a0038933

Reeve, H. K. & Sherman, P. W. (1993). Adaptation and the goals of evolutionary research. *The Quarterly Review of Biology*, *68*(1), 1–32. doi:10.1086/417909

Reimer, N. K., Becker, J. C., Benz, A., Christ, O., Dhont, K., Klocke, U., … Hewstone, M. (2017). Intergroup contact and social change: Implications of negative and positive contact for collective action in advantaged and disadvantaged groups. *Personality and Social Psychology Bulletin*, *43*(1), 121–136. doi:10.1177/0146167216676478

Ritchie, S. J. (2015). *Intelligence: All that matters*. London, UK: Hodder & Stoughton.

Ritter, S. M., Damian, R. I., Simonton, D. K., van Baaren, R. B., Strick, M., Derks, J. & Dijksterhuis, A. (2012). Diversifying experiences enhance cognitive flexibility. *Journal of Experimental Social Psychology*, *48*(4), 961–964. doi:10.1016/j.jesp.2012.02.009

Ritter, S. M., Kühn, S., Müller, B. C. N., van Baaren, R. B., Brass, M. & Dijksterhuis, A. (2014). The creative brain: Corepresenting schema violations enhances TPJ activity and boosts cognitive flexibility. *Creativity Research Journal*, *26*(2), 144–150. doi:10.1080/10400419.2014.901061

Roccas, S. & Brewer, M. B. (2002). Social identity complexity. *Personality and Social Psychology Review*, *6*(2), 88–106. doi:10.1207/s15327957pspr0602_01

Roets, A. & Van Hiel, A. (2011). Allport's prejudiced personality today: Need for closure as the motivated cognitive basis of prejudice. *Current Directions in Psychological Science*, *20*(6), 349–354. doi:10.1177/0963721411424894

Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T. & Lyon, D. (1989). Evidence for terror management theory: I. the effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of Personality and Social Psychology*, *57*(4), 681–690. doi:10.1037/0022-3514.57.4.681

Rosseel, Y. (2012). lavaan: An R package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. doi:10.18637/jss.v048.i02

Rothbart, M. & John, O. P. (1985). Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact. *Journal of Social Issues*, *41*(3), 81–104. doi:10.1111/j.1540-4560.1985.tb01130.x

Routledge, C. & Juhl, J. (2012). The creative spark of death: The effects of mortality salience and personal need for structure on creativity. *Motivation and Emotion*, *36*(4), 478–482. doi:10.1007/s11031-011-9274-1

Rubin, M. (2017). Towards a multiple motives meta-theory for social psychology. *Journal of Theoretical Social Psychology*, *1*(1), 15–20. doi:10.1002/jts5.3

Rubin, M., Paolini, S. & Crisp, R. J. (2010). A processing fluency explanation of bias against migrants. *Journal of Experimental Social Psychology*, *46*(1), 21–28. doi:10.1016/j.jesp.2009.09.006

Rubin, M., Paolini, S. & Crisp, R. J. (2011). The relationship between the need for closure and deviant bias: An investigation of generality and process. *International Journal of Psychology*, *46*(3), 206–213. doi:10.1080/00207594.2010.537660

Rubin, M., Paolini, S. & Crisp, R. J. (2013). Linguistic description moderates the evaluations of counterstereotypical people. *Social Psychology*, *44*(4), 289–298. doi:10.1027/1864-9335/a000114

Ryan, R. & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation. *American Psychologist*, *55*(1), 68–78. doi:10.1037/0003-066X.55.1.68

Saguy, T., Tausch, N., Dovidio, J. F. & Pratto, F. (2009). The irony of harmony: Intergroup contact can produce false expectations for equality. *Psychological Science*, *20*(1), 114–121. doi:10.1111/j.1467-9280.2008.02261.x

Sandberg, S. (2013). *Lean in: Women, work, and the will to lead*. London, UK: WH Allen.

Sassenberg, K. & Moskowitz, G. B. (2005). Don't stereotype, think different! Overcoming automatic stereotype activation by mindset priming. *Journal of Experimental Social Psychology*, *41*(5), 506–514. doi:10.1016/j.jesp.2004.10.002

Schmid, K., Hewstone, M. & Tausch, N. (2014). Secondary transfer effects of intergroup contact via social identity complexity. *British Journal of Social Psychology*, *53*(3), 443–462. doi:10.1111/bjso.12045

Schmitt, D. P. & Pilcher, J. J. (2004). Evidence of evaluating psychological adaptation: How do we know one when we see one? *Psychological Science*, *15*(10), 643–649. doi:https://doi.org/10.1111/j.0956-7976.2004.00734.x

Schulz-Hardt, S., Frey, D., Lüthgens, C. & Moscovici, S. (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology*, *78*(4), 655–669. doi:10.1037OTO22-3514.78.4.655

Schützwohl, A. & Borgstedt, K. (2005). The processing of affectively valenced stimuli: The role of surprise. *Cognition and Emotion*, *19*(4), 583–600. doi:10.1080/02699930441000337

Schützwohl, A. & Reisenzein, R. (1999). Children's and adults' reactions to a schema- discrepant event: A developmental analysis of surprise. *Interna-*

tional Journal of Behavioural Development, *23*(1), 37–62. doi:10 . 1080 / 016502599383991

Shafir, E. & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology*, *24*(4), 449–474. doi:10.1016/0010-0285(92)90015-T

Shah, A. K., Mullainathan, S. & Shafir, E. (2012). Some consequences of having too little. *Science*, *338*(6107), 682–685. doi:10.1126/science.1222426

Sheldon, K. M., Elliot, A. J., Kim, Y. & Kasser, T. (2001). What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of Personality and Social Psychology*, *80*(2), 325–339. doi:10.1037/0022-3514.80.2.325

Sherif, M., White, J., Sherif, C. W., Hood, W. R. & Harvey, O. J. (1961). *Intergroup conflict and cooperation: The Robbers' Cave experiment.* Norman, OK: Oklahoma Book Exchange.

Shrout, P. E. & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, *69*(1), 487–510. doi:10.1146/annurev-psych-122216-011845

Silvia, P. J. (2008). Appraisal components and emotion traits: Examining the appraisal basis of trait curiosity. *Cognition and Emotion*, *22*(1), 94–113. doi:10.1080/02699930701298481

Silvia, P. J., Martin, C. & Nusbaum, E. C. (2009). A snapshot of creativity: Evaluating a quick and simple method for assessing divergent thinking. *Thinking Skills and Creativity*, *4*(2), 79–85. doi:10.1016/j.tsc.2009.06.005

Simmons, J. P., Nelson, L. D. & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. doi:10.1177/0956797611417632

Simon, B. & Pettigrew, T. F. (1990). Social identity and perceived group homogeneity: Evidence for the ingroup homogeneity effect. *European Journal of Social Psychology*, *20*(4), 269–286. doi:10.1002/ejsp.2420200402

Simonton, D. K. (1997). Foreign influence and national achievement: The impact of open milieus on Japanese civilization. *Journal of Personality and Social Psychology*, *72*(1), 86–94. doi:10.1037/0022-3514.72.1.86

Simonton, D. K. (2000). Creativity: Cognitive, personal, developmental, and social aspects. *American Psychologist*, *55*(4), 407–421. doi:10.1037//0003

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York, NY: Appleton-Century.

Sokolov, E. N., Spinks, J. A., Näätänen, R. & Lyytinen, H. (2002). *The orienting response in information processing*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Sommers, S. R. (2006). On racial diversity and group decision making: Identifying multiple effects of racial composition on jury deliberations. *Journal of Personality and Social Psychology*, *90*(4), 597–612. doi:10.1037/0022-3514.90.4.597

Sport England. (2014). This Girl Can. Retrieved April 23, 2018, from https://www.sportengland.org/news-and-features/news/2014/october/31/this-girl-can/

Sport England. (2016). Record number of women get active. Retrieved April 23, 2018, from https://www.sportengland.org/news-and-features/news/2016/december/8/record-numbers-of-women-getting-active/

Sport England. (2017). Active People Survey. Retrieved April 23, 2018, from https://www.sportengland.org/research/about-our-research/active-people-survey/

Srull, T. K. & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review*, *96*(1), 58–83. doi:10.1037/0033-295X.96.1.58

Stanovich, K. E. & West, R. F. (2002). Individual differences in reasoning: Implications for the rationality debate? In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 421–440). New York, NY: Cambridge University Press. doi:http://dx.doi.org/10.1017/CBO9780511808098.026

Stanovich, K. E. & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*(4), 672–695. doi:10.1037/0022-3514.94.4.672

Stathi, S. & Crisp, R. J. (2010). Intergroup contact and the projection of positivity. *International Journal of Intercultural Relations*, *34*(6), 580–591. doi:10.1016/j.ijintrel.2010.04.002

Stathi, S., Crisp, R. J. & Hogg, M. A. (2011). Imagining intergroup contact enables member-to-group generalization. *Group Dynamics: Theory, Research, and Practice*, *15*(3), 275–284. doi:10.1037/a0023752

Steele, C. M. & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, *45*(1), 5–19. doi:10.1037/0022-3514.45.1.5

Stevens, L. E. & Fiske, S. T. (1995). Motivation and cognition in social life: A social survival perspective. *Social Cognition*, *13*(3), 189–214. doi:10.1521/soco.1995.13.3.189

Strack, F. & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247. doi:10.1207/s15327957pspr0803_1

Tadmor, C. T., Hong, Y.-y., Chao, M. M., Wiruchnipawan, F. & Wang, W. (2012). Multicultural experiences reduce intergroup bias through epistemic unfreezing. *Journal of Personality and Social Psychology*, *103*(5), 750–772. doi:10.1037/a0029719

Tadmor, C. T., Tetlock, P. E. & Peng, K. (2009). Acculturation strategies and integrative complexity. *Journal of Cross-Cultural Psychology*, *40*(1), 105–139. doi:10.1177/0022022108326279

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. doi:10.1002/ejsp.2420010202

Tajfel, H. & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks-Cole.

Tashakkori, A. & Teddlie, C. (2010). *Sage handbook of mixed methods in social & behavioral research* (1st ed.). Thousand Oaks, CA: SAGE Publications.

Terman, L. M. & Merrill, M. A. (1972). *Stanford Binet intelligence scale.* (3rd ed.). Houghton Mifflin Harcourt.

Tesser, A. (2000). On the confluence of self-esteem maintenance mechanisms. *Personality and Social Psychology Review*, *4*(4), 290–299. doi:10.1207/S15327957PSPR0404_1

Tesser, A. & Cornell, D. P. (1991). On the confluence of self processes. *Journal of Experimental Social Psychology*, *27*(6), 501–526. doi:10.1016/0022-1031(91)90023-Y

Tesser, A. (2001). On the plasticity of self-defense. *Current Directions in Psychological Science*, *10*(2), 66–69. doi:10.1111/1467-8721.00117

The Cochrane Collaboration. (2014). Review Manager (RevMan). Copenhagen: The Nordic Cochrane Centre. Retrieved from http://www.cochrane.org/about-us/citing-our-products

The New York Times. (2016). Prince: In memoriam. Retrieved April 23, 2018, from https://twitter.com/nytimes/status/723233362228465664

Thompson, E. P., Roman, R. J., Moskowitz, G. B., Chaiken, S. & al, E. (1994). Accuracy motivation attenuates covert priming: The systematic reprocessing of social information. *Journal of Personality and Social Psychology*, *66*(3), 474–489. doi:10.1037/0022-3514.66.3.474

Thompson, M. M., Zanna, M. P. & Griffin, D. W. (1995). Let's not be indifferent about (attitudinal) ambivalence. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (Chap. 14, pp. 361–386). New York, NY: Psychology Press.

Tomaka, J., Blascovich, J., Kelsey, R. M. & Leitten, C. L. (1993). Subjective, physiological, and behavioral effects of threat and challenge appraisal. *Journal*

of Personality and Social Psychology, *65*(2), 248–260. doi:10.1037/0022-3514.65.2.248

Tomaka, J., Blascovich, J., Kibler, J. & Ernst, J. M. (1997). Cognitive and physiological antecedents of threat and challenge appraisal. *Journal of Personality and Social Psychology*, *73*(1), 63–72. doi:10.1037/0022-3514.73.1.63

Toplak, M. E., West, R. F. & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking and Reasoning*, *20*(2), 147–168. doi:10.1080/13546783.2013.844729

Topolinski, S. & Strack, F. (2015). Corrugator activity confirms immediate negative affect in surprise. *Frontiers in Psychology*, *6*(2), 1–8. doi:10.3389/fpsyg.2015.00134

Trope, Y. & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463. doi:10.1037/a0018963

Turner, R. N., Crisp, R. J. & Lambert, E. (2007). Imagining intergroup contact can improve intergroup attitudes. *Group Processes & Intergroup Relations*, *10*(4), 427–441. doi:10.1177/1368430207081533

Turner, R. N., West, K. & Christie, Z. (2013). Out-group trust, intergroup anxiety, and out-group attitude as mediators of the effect of imagined intergroup contact on intergroup behavioral tendencies. *Journal of Applied Social Psychology*, *43*(S2), 196–205. doi:10.1111/jasp.12019

Tversky, A. & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, *106*(4), 1039–1061. doi:10.2307/2937956

van der Meer, T. & Tolsma, J. (2014). Ethnic diversity and its effects on social cohesion. *Annual Review of Sociology*, *40*, 459–478. doi:10.1146/annurev-soc-071913-043309

van Egmond, M. C., Kühnen, U. & Li, J. (2013). Mind and virtue: The meaning of learning, a matter of culture? *Learning, Culture and Social Interaction*, *2*(3), 208–216. doi:10.1016/j.lcsi.2013.06.002

van Harreveld, F., Nohlen, H. U. & Schneider, I. K. (2015). The ABC of ambivalence: Affective, behavioral, and cognitive consequences of attitudinal conflict. *Advances in Experimental Social Psychology*, *52*, 285–324. doi:10.1016/bs.aesp.2015.01.002

van Tilburg, W. A. & Igou, E. R. (2012). On boredom: Lack of challenge and meaning as distinct boredom experiences. *Motivation and Emotion*, *36*(2), 181–194. doi:10.1007/s11031-011-9234-9

Vasiljevic, M. & Crisp, R. J. (2013). Tolerance by surprise: Evidence for a generalized reduction in prejudice and increased egalitarianism through novel category combination. *PLoS ONE*, *8*(3), 1–9. doi:https://doi.org/10.1371/journal.pone.0057106

Vess, M., Routledge, C., Landau, M. J. & Arndt, J. (2009). The dynamics of death and meaning: The effects of death-relevant cognitions and personal need for structure on perceptions of meaning in life. *Journal of Personality and Social Psychology*, *97*(4), 728–744. doi:10.1037/a0016417

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Vincent, L. C., Emich, K. J. & Goncalo, J. A. (2013). Stretching the moral gray zone. *Psychological Science*, *24*(4), 595–599. doi:10.1177/0956797612458806

Visintin, E. P., Birtel, M. D. & Crisp, R. J. (2017). The role of multicultural and colorblind ideologies and typicality in imagined contact interventions. *International Journal of Intercultural Relations*, *59*(4), 1–8. doi:10.1016/j.ijintrel.2017.04.010

Voci, A. & Hewstone, M. (2003). Intergroup contact and prejudice toward immigrants in Italy: The mediational role of anxiety and the moderational role of group salience. *Group Processes & Intergroup Relations*, *6*(1), 37–54. doi:10.1177/1368430203006001011

von Stumm, S., Hell, B. & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science*, *6*(6), 574–588. doi:10.1177/1745691611421204

Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, *23*(1), 73–82. doi:10.1177/0963721413512856

Webster, D. M. & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, *67*(6), 1049–1062. doi:10.1037/0022-3514.67.6.1049

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore, MD: Williams & Wilkins Co. doi:10.1037/10020-000

Werner, K. M., Milyavskaya, M., Foxen-Craft, E. & Koestner, R. (2016). Some goals just feel easier: Self-concordance leads to goal progress through subjective ease,

not effort. *Personality and Individual Differences*, *96*, 237–242. doi:10.1016/j.paid.2016.03.002

West, R. F., Toplak, M. E. & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, *100*(4), 930–941. doi:10.1037/a0012842

Wickham, H. (2009). *Ggplot2: Elegant graphics for data analysis*. New York, NY: Springer-Verlag. Retrieved from http://ggplot2.org

Widfeldt, A. & Brandenburg, H. (2017). What kind of party is the UK independence party? The future of the extreme right in Britain or just another Tory party? *Political Studies*, 1–24. doi:10.1177/0032321717723509

Wilder, D. A. (1984). Intergroup contact: The typical member and the exception to the rule. *Journal of Experimental Social Psychology*, *20*(2), 177–194. doi:https://doi.org/10.1016/0022-1031(84)90019-2

Wilder, D. A., Simon, A. F. & Faith, M. (1996). Enhancing the impact of counterstereotypic information: Dispositional attributions for deviance. *Journal of Personality & Social Psychology*, *71*(2), 276–287. doi:10.1037/0022-3514.71.2.276

Wilder, D. A. & Thompson, J. E. (1980). Intergroup contact with independent manipulations on in-group and out-group interaction. *Journal of Personality and Social Psychology*, *38*(4), 589–603. doi:10.1037//0022-3514.38.4.589

Williams, G. C. (1996). *Adaptation and natural selection*. Princeton, NJ: Princeton University Press.

Wilson, G. D. (1973). A dynamic theory of conservatism. In G. D. Wilson (Ed.), *The psychology of conservatism.* Oxford, England: Academic Press.

Wohl, M. J. & Branscombe, N. R. (2005). Forgiveness and collective guilt assignment to historical perpetrator groups depend on level of social category inclusiveness. *Journal of Personality and Social Psychology*, *88*(2), 288–303. doi:10.1037/0022-3514.88.2.288

Woodworth, R. S. & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions (I). *Psychological Review*, *8*(3), 247–261. doi:10.1037/h0074898

Wright, S. C., Aron, A., McLaughlin-Volpe, T. & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, *73*(1), 73–90. doi:10.1037/0022-3514.73.1.73

Zajonc, R. B. (1998). Emotions. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 591–632). New York, NY: McGraw-Hill.

Zenith Media. (2017). 26% of media consumption will be mobile in 2019. Retrieved April 27, 2018, from https://www.zenithmedia.com/26-of-media-consumption-will-be-mobile-in-2019/