INFLUENCE OF INTERNAL VERSUS EXTERNAL EGALITARIAN GOALS ON STEREOTYPE ACCESSIBILITY FOLLOWING GOAL–BEHAVIOUR DISCREPANCIES

by

LAURA FRANCES KIMBERLEY

(B.Sc. University of Birmingham)

A thesis submitted to the University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology
College of Life and Environmental Sciences
The University of Birmingham
October 2014

ABSTRACT

The mechanism through which goals influence stereotype activation/control is unclear. This thesis aimed to shed some light on the mechanism through which internally-generated and externally-imposed goals influence stereotype activation by applying Fishbach and colleagues' model of goal progress. Across six experiments, I demonstrated that (1) internal, but not external, egalitarian goals result in less stereotype accessibility (consistent with egalitarian goal pursuit) following a potentially large goal–behaviour discrepancy (Experiment 2), (2) external, but not internal, egalitarian goals may result in less stereotype accessibility (consistent with egalitarian goal pursuit) when contemplating past success (Experiment 1), but not following a potentially small goal–behaviour discrepancy or a potential goal–behaviour match (Experiments 3–5), and (3) internal egalitarian goals result in a motivation to demonstrate inclusiveness by including racially ambiguous faces in the ingroup, but only following a potential goal–behaviour discrepancy. External egalitarian goals irrespective of goal–behaviour discrepancy size, and internal egalitarian goals following a potential goal–behaviour match, both result in a motivation to be accurate when categorising racially ambiguous faces (Experiment 6). These findings suggest that the source of a goal (i.e., internal vs. external), and discrepancy size for internal, but not external, goals, are important parts of the mechanism through which goals influence social categorisation and stereotype activation.

ACKNOWLEDGEMENTS

The completion of this thesis would not have been possible without the help of a number of people. First, I would like to thank my two supervisors: Dr Brandon Stewart and Dr Kimberly Quinn. I would like to thank Brandon for inspiring me to read widely, and for teaching me to think critically about my research ideas, both of which have made me a better academic; and Kim for encouraging me to think outside of the box when developing ideas, and for her belief in my academic ability, both of which have been a true source of motivation. I appreciate the time and effort that they have both given over the past three years, and I have enjoyed working with them.

Second, I would also like to thank the School of Psychology at the University of Birmingham for awarding me a studentship to complete my PhD research. I would like to thank the studentship committee members for choosing to support this research.

Finally, I would like to thank my friends and my family. I would like to thank Chelsea Slater for her friendship which has been a true source of support throughout the past three years. I would also like to thank my family for their understanding and support while completing my PhD. I would particularly like to thank my mum for the work ethic that she taught me. I couldn't have completed this thesis without it. I would also like to thank my nan for her unwavering belief that I can do anything I put my mind to. Last, but not least, I would like to thank Barnaby Smith for his love and support, for keeping me sane, and for listening to me think out loud last thing at night.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1

THE IMPACT OF INTERNAL AND EXTERNAL EGALITARIAN GOALS ON STEREOTYPE ACTIVATION

Most societies value acting egalitarian towards stigmatised groups (Fehr & Sassenberg, 2010). Many people within society personally adopt an egalitarian goal because they believe in and value acting fairly and tolerantly of others (internally-generated egalitarian goal; Plant & Devine, 1998). Others, however, adopt an egalitarian goal to avoid social sanctions that arise from failing to comply with external pressure to act egalitarian (externally-imposed egalitarian goal; Plant & Devine, 1998). Research has demonstrated that individuals who are internally motivated to act egalitarian exhibit less stereotype endorsement and prejudice than individuals who are externally motivated to act egalitarian (e.g., Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Plant & Devine, 1998). However, the mechanism through which internally-generated and externally-imposed goals influence stereotype activation is not fully understood. The present thesis aimed to apply Fishbach and colleagues' model (Fishbach & Dhar, 2005, 2007; Fishbach, Zhang, & Koo, 2009; Koo & Fishbach, 2008) of goal progress and motivation to stereotype activation. The model asserts that motivation and goal adherence are a function of the discrepancy between actual and desired goal states, and whether movement towards a goal is construed in terms of progress or commitment to the goal. The application of Fishbach and colleagues' model may shed some light on the mechanism through which internally-generated and externally-imposed goals influence stereotype activation.

1.0 General Introduction

Why does stereotyping and prejudice prevail despite a pervasive egalitarian social norm and the general desire to maintain an egalitarian self-image (see Fehr & Sassenberg, 2010; Fehr, Sassenberg, & Jonas, 2012; Gaertner & Dovidio, 1986; McConahay, 1986; Monteith, Deneen, & Tooman, 1996)? According to Fehr and colleagues (Fehr & Sassenberg, 2010; Fehr et al., 2012), controlling stereotyping and prejudice is not always possible. When a person is encountered, the perceptual properties of the face (e.g., dark skin) are often automatically processed, which subsequently signals the category membership of the person (e.g., Black) and activates associated stereotypes and/or prejudice (e.g., aggressive; Allport, 1954; Bargh, 1999; Brewer, 1988; Devine, 1989; Fiske & Neuberg, 1990; Livingston & Brewer, 2002; Tajfel, 1969). Although this process is not inevitable (see Blair, 2002, for a review), the ease with which stereotypes and prejudice can be automatically activated by cues in the environment (e.g., Black athlete) or external influences (e.g., racist jokes) may explain why stereotyping and prejudice remain pervasive despite people's best intentions to act egalitarian (Fehr et al., 2012). Therefore, a key issue relates to when people are more or less likely to avoid using stereotypes and prejudice.

The present research aimed to determine the role of goal source (internal vs. external egalitarian goal) and goal–behaviour discrepancies (distance between actual behaviour and standards for behaviour) in determining stereotype activation. In the present chapter, I will begin by defining the key terms used throughout the present thesis. Then, I will discuss the automatic nature of stereotype activation. Next, I will discuss the stereotype-control strategies that may aid people in controlling stereotypes. Then, I will focus on goals in more detail, discussing the influence of both internal and external motivation on stereotype activation, and considering the role of goal–behaviour discrepancies. Finally, I will present a self-regulation

model developed by Fishbach and colleagues (Fishbach & Dhar, 2005, 2007; Fishbach et al., 2009; Koo & Fishbach, 2008), which may further our understanding of how goal source and goal–behaviour discrepancies influence stereotype activation.

1.1 Terminology

1.1.1 Stereotype. Similar to Mokowitz (2010), I define a stereotype as consisting of both a cognitive and a motivational component. Cognitively, a stereotype is the mental representation of a particular group that consists of the perceiver's knowledge, beliefs, and expectations about the group (Hamilton & Trolier, 1986; Moskowitz, 2010; Moskowitz & Li, 2011). A stereotype may become cognitively accessible within a perceiver's mind such that the stereotype is ready to be used in subsequent judgements (*stereotype activation;* Kunda & Spencer, 2003). Once a stereotype is cognitively accessible, the stereotype *may* then be used to judge a member of the group to which the stereotype pertains (*stereotype application/use*; Kunda & Spencer, 2003).

Motivationally, a stereotype is a "cognitive tool" that serves goals (Gilbert & Hixon, 1991; Moskowitz, 2010; Moskowitz & Li, 2011). As a result, some goals may aid stereotype activation and application, and other goals may inhibit stereotype activation and application (see Kunda & Spencer (2003) and Moskowitz (2010), for reviews). For example, stereotypes may aid comprehension goals by reducing the wealth of social information available to the social perceiver; hence, such goals may foster stereotype activation and application (Macrae, Bodenhausen, Milne, & Jetten, 1994; Pendry & Macrae, 1996; Kunda & Spencer, 2003). Conversely, stereotypes may also disrupt goals to avoid prejudice or goals to act egalitarian; hence, such goals may foster the inhibition of stereotype activation and application (Moskowitz, 2010; Kunda & Spencer, 2003).

1.1.2 Prejudice. I adopt the widely used definition of prejudice within the field: "[a] negative attitude toward an out-group" (Wittenbrink, Judd, & Park, 1997, p. 265; see also, Allport, 1954; Ashmore, 1970; Dovidio & Gaertner, 1986; Esses, Haddock & Zanna, 1993; Stephan, 1985; Stroebe & Insko, 1989).

1.1.3 Egalitarian goal. Based on Kruglanski, Shah, Fishbach, Friendman, Chun, and Sleeth-Keppler's (2002) theory of goal systems, I define a goal as "striving to attain specific desirable objectives" (Kruglanski et al., 2002, p. 9). Fundamentally, a goal is a discrepancy between the perceiver's current state and the desired end-state (Moskowitz, Li, Ignarri, & Stone, 2011). Furthermore, perceiving a discrepancy results in a tension-state that compels the perceiver to act consistently with the goal in order to reduce the size of the discrepancy and alleviate the tension state (Moskowitz & Li, 2011; Moskowitz et al., 2011).

An egalitarian goal, then, constitutes striving to "act fair, just, and tolerant of others as well as treating people equally regardless of whether they differ from you, and regardless of their ethnicity, religious background, gender, sexual orientation, physical appearance, etc." (Moskowitz, 2002, p. 401). A discrepancy between a perceiver's current egalitarian behaviour (i.e., the current state) and their standard of egalitarian behaviour (i.e., the desired end state) results in a tension state that drives a perceiver to act egalitarian to reduce the discrepancy and alleviate the tension state (Moskowitz & Li, 2011; Moskowitz et al., 2011). A perceiver may act consistently with an egalitarian goal by consciously suppressing stereotypes in situations where the likelihood of stereotype activation occurring is high, such as after hearing a joke that relies on stereotypes (Moskowitz et al., 2011; Fehr et al., 2012). Perceivers may also act consistently with an egalitarian goal by nonconsciously selectively attending to goal-relevant stimuli (e.g., a Black face) which provide opportunities to respond consistently with the goal (Moskowitz et al., 2011).

A brief discussion of whether goals are analogous to values is relevant here. Rokeach (1973) distinguished between two types of values: instrumental and terminal values. Instrumental values reflect behavioural methods (i.e., mode of conduct) of achieving terminal values. Examples of instrumental values include broad-mindedness, love, courage, and honesty. Terminal values reflect a perceiver's life-long objectives (i.e., end-state of existence). Examples of terminal values include happiness, equality, pleasure, and wisdom. I believe that an egalitarian goal may be the concrete manifestation of holding a broad-minded instrumental value that serves an equality terminal value. Specifically, a discrepancy between one's current egalitarian behaviour and one's standard of egalitarian behaviour motivates a perceiver to act consistently with their egalitarian goal by controlling one's stereotyping and prejudice which in turn serves the perceiver's values.

1.2 The Automaticity of Stereotype Activation

Debate about the automaticity of stereotype activation has a long history within social psychology. For example, Allport (1954) argued that categorising a person according to the categories to which they belong (e.g., age, gender, and ethnicity) is an unavoidable process. In addition, categorising a person often leads to the activation of relevant stereotypical information associated with those categories (e.g., dark skin automatically activates the stereotype of Black people; Allport, 1954; Brewer, 1988; Devine, 1989; Fiske & Neuberg, 1990; Tajfel, 1969). Many theorists argue that stereotype activation is beneficial for the stereotype user because the amount of information that needs to be processed during social interaction is reduced (i.e., stereotypes are a mental shortcut) and this frees up sparse cognitive resources for other tasks (e.g., Lippman, 1922; Macrae, Milne, & Bodenhausen, 1994; Macrae, Stanger, & Milne, 1994; Sherman, Lee, Bessenoff, & Frost, 1998). The repeated and consistent use of stereotypes as a mental short cut from an early age creates an

automatic link between the person in the environment and relevant stereotypes (see Bargh & Chartrand, 1999; Devine, 1989). Consequently, for many decades, researchers assumed that stereotype activation automatically and inevitably follows person perception (e.g., Allport, 1954; Brewer, 1988; Devine, 1989; Fiske & Neuberg, 1990; Tajfel, 1969).

Early theoretical and empirical research supported the assumption that stereotype activation is automatic and inevitable. Although people often moderate their *explicit* expressions of stereotypes and prejudice, they still tend to exhibit *implicit* stereotype activation and prejudice (e.g., Cunningham, Preacher, & Banaji, 2001; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998). For example, Dovidio et al. (1997, Experiment 1) found that although participants exhibited low levels of prejudice towards Black people on explicit questionnaire measures, participants nonetheless exhibited prejudice towards Black people on an implicit measure.

Devine (1989) developed the dissociation model to explain the disparity between implicit and explicit measures of stereotype activation and prejudice. The model proposes that stereotypes are automatically activated by both high- and low-prejudiced people due to being learned at an early age. Although low-prejudiced people hold personal beliefs that contradict stereotypes, these personal beliefs are not as accessible as stereotypes because they developed later. As a result, both high- and low-prejudiced people exhibit prejudice on implicit measures because conscious control is prevented. In contrast, only high-prejudiced participants exhibit prejudice on explicit measures because low-prejudiced people consciously inhibit stereotypes and respond according to their personal beliefs instead. Devine's (1989) dissociation model indicates that although people may inhibit the cultural stereotype and respond according to

their personal beliefs at a conscious level, the cultural stereotype will remain the automatic response (except see Lepore & Brown, 1997).

Although stereotype activation can clearly operate automatically without conscious awareness, Gilbert and Hixon (1991) argued that stereotype activation need not be unconditionally automatic. In their first experiment, participants completed more word fragments in an Asian-stereotype-congruent manner when the experimenter was Asian compared to White. However, this difference disappeared when participants completed a cognitively-demanding rehearsal task while completing the word fragments. Gilbert and Hixon concluded that the cognitively-demanding rehearsal task prevented information associated with the category (e.g., stereotypes) from being activated, and thus that stereotype activation is conditionally rather than unconditionally automatic.

However, subsequent research by Spencer, Fein, Wolfe, Fong, and Duinn (1998) demonstrated that participants who were motivated to use stereotypes in order to restore their sense of self-worth exhibited stereotype activation even when under cognitive load. This finding has implications for Gilbert and Hixon's argument that cognitive resources play a key role in determining whether stereotype activation occurs (Bodenhausen, Todd, & Richeson, 2009). Specifically, Bodenhausen et al. (2009) argued that if the participants who completed the cognitively-demanding rehearsal task in Gilbert and Hixon's experiment had been motivated to think about the experimenter, they may have exhibited stereotype activation despite low cognitive resources. Therefore, motivation may be a more important determinant of whether stereotype activation occurs than the availability of cognitive resources (Bodenhausen et al., 2009).

Further research has identified a host of additional moderators of stereotype activation and prejudice (see Blair (2002) and Quinn, Macrae, & Bodenhausen (2003) for reviews).

Some of the moderators relate to the individual, including processing objectives/goals (e.g., Galinsky & Moskowitz, 2000; Macrae, Bodenhausen, Milne, Thorn, & Castelli, 1997; Stewart & Payne, 2008; Wheeler & Fiske, 2005), expectations (e.g., Blair & Banaji, 1996), chronic motivation (e.g., Moskowitz, Gollwitzer, Wasel, & Schaal, 1999; Moskowitz, Salomon, & Taylor, 2000) and prejudice-level (e.g., Lepore & Brown, 1997). Some of the moderators relate to the situation, including contextual information (e.g., Barden, Maddux, Petty, & Brewer, 2004; Dasgupta & Greenwald, 2001; Lowery, Hardin, & Sinclair, 2001; Wittenbrink, Judd, & Park, 2001a), and task differences (e.g., Wittenbrink, Judd, & Park, 2001b).

In sum, it is clear from empirical research that stereotype activation is not unconditionally automatic. Yet, cues in the environment (e.g., a Black athlete) or external influences (e.g., racist jokes) may activate stereotypes despite the social perceivers best intentions to act egalitarian (Fehr et al., 2012). As a result, researchers have examined a number of stereotype-control strategies to help people avoid the use of stereotypes (see Blair, 2002, for a review).

1.3 Reactive Stereotype-Control Strategies

Most stereotype-control strategies are reactive in nature. Once a person consciously detects the activation of stereotypes, strategies to prevent stereotypes from influencing behaviour are implemented (Moskowitz, 2010; Moskowitz & Li, 2011; Moskowitz & Stone, 2012). Three well-known reactive stereotype-control strategies are stereotype suppression, counter-stereotypes, and training (see Blair, 2002, for these and other stereotype-control strategies).

1.3.1 Stereotype suppression. Several researchers have demonstrated that stereotype suppression results in less stereotype use (e.g., Galinsky & Moskowitz, 2000; Liberman & Förster, 2000; Macrae, Bodenhausen, Milne, & Jetten, 1994; see Monteith, Sherman, &

Devine, 1998, for a review). For example, Macrae, Bodenhausen, Milne, and Jetten (1994, Experiment 1) asked participants to write an essay about a typical day in the life of a skinhead. Before writing the essay, half of the participants were instructed to suppress their stereotypes. The results demonstrated that participants instructed to suppress their stereotypes wrote essays that were significantly less stereotypical of skinheads than participants who received no suppression instructions.

However, successfully suppressing stereotypes on one task can have the ironic effect of heightening stereotype activation and application on a subsequent task, known as the stereotype rebound effect (Galinsky & Moskowitz, 2000, 2007; Liberman & Förster, 2000; Macrae, Bodenhausen, Milne, & Jetten, 1994; see Monteith, Sherman, & Devine, 1998, for a review). For example, Macrae, Bodenhausen, Milne, and Jetten (1994) showed that after participants successfully suppressed stereotypes while writing an essay about a typical day in the life of a skinhead, they subsequently sat further from a skinhead (Experiment 2), and exhibited greater activation of the skinhead stereotype (indexed by faster reaction times to skinhead-stereotypic than stereotype-neutral words; Experiment 3), than participants who received no suppression instructions. Macrae, Bodenhausen, Milne, and Jetten proposed the following mechanism to explain the stereotype rebound effect: The act of stereotype suppression involves monitoring consciousness for any trace of stereotypes which inadvertently increases the accessibility of stereotypes. The heightened accessibility of stereotypes takes time to dissipate and will likely affect subsequent information processing (see Wyer, Sherman, & Stroessner (2000) and Gordijn, Hindriks, Koomen, Dijksterhuis, & Van Knippenberg (2004), for similar arguments). Nevertheless, subsequent research has identified a series of moderators that circumvent the stereotype rebound effect, including

motivation (e.g., Gordijn et al., 2004; Wyer et al., 2000; Wyer, 2007) and prejudice-level (e.g., Monteith, Spicer, & Tooman, 1998).

1.3.2 Counter-stereotypes. Actively thinking about counter-stereotypes has been shown to reduce stereotype activation. For example, Blair and Banaji (1996) investigated the effect of counter-stereotype expectancies on the activation of gender stereotypes. Stereotype activation was assessed using a sequential priming task; participants were presented with a male stereotype-congruent word (e.g., *ambitious*) or female stereotype-congruent word (e.g., *perfume*), and then categorised a forename as male or female. On half of the trials the words and names formed stereotype-consistent pairs (e.g., *gentle–Jane* and *strong–John*), and on the other half of the trials the words and names formed counter-stereotype pairs (e.g., *strong–Jane* and *gentle–John*). Greater activation of gender stereotypes is indicated when participants respond more quickly and more accurately to stereotype-consistent pairs compared to counter-stereotype pairs. Blair and Banaji (1996) found that participants instructed to expect counter-stereotype pairs exhibited less activation of gender stereotypes than participants instructed to expect stereotype-consistent pairs.

Additionally, actively thinking about counter-stereotype exemplars has been shown to reduce stereotype activation. For example, Blair, Ma, and Lenton (2001, Experiments 1–3) investigated the effect of imagining a counter-stereotypic female exemplar on the activation of the female stereotype. Stereotype activation was assessed using an Implicit Association Test (IAT; Greenwald et al., 1998), which assessed the strength of association between gender and physical strength. Blair et al. found that participants who imagined a counter-stereotypic woman (i.e., a strong woman) prior to completing the IAT exhibited less activation of the female stereotype than female participants who imagined a stereotypical woman (i.e., a weak

woman), a neutral event (i.e., a vacation), or did not engage in any mental imagery prior to completing the IAT.

1.3.3 Training. With respect to training, Kawakami, Dovidio, Mill, Hermsen, and Russin (2000) demonstrated that after training participants to negate stereotypes, stereotype activation was reduced. In their second experiment, for example, participants in the experimental condition completed skinhead stereotype negation training; participants pressed "no" in response to information congruent with the skinhead stereotype, and "yes" to information incongruent with the skinhead stereotype (in this case, information congruent with the elderly stereotype). After the training, these participants completed a primed Stroop task; activation of the skinhead stereotype was indexed by longer colour naming times for skinhead-stereotype-congruent words following a "skinhead" prime versus an "elderly" prime. In contrast, participants in the control condition only completed the primed Stroop task. Analysis revealed that participants who completed the skinhead stereotype negation training exhibited less activation of the skinhead stereotype than participants in the control condition, both immediately after training and after 24 hours had passed. The successful role of training in the reduction of stereotype activation and prejudice has been supported in subsequent research (e.g., Kawakami, Dovidio, & van Kamp, 2007; Kawakami, Phills, Steele, & Dovidio, 2007; Plant, Peruche, & Butz, 2005; Peruche & Plant, 2006; Rudman, Ashmore, & Gary, 2001).

1.4 Proactive Stereotype-Control Strategies

In addition to the reactive stereotype-control strategies discussed above, stereotypes can also be controlled proactively, through the goals that we hold during social interaction (see Kunda & Spencer, 2003, for a review). Goals are a proactive stereotype-control strategy because the conscious detection of stereotype activation is not required for the strategy to be

implemented, unlike reactive stereotype-control strategies (Moskowitz, 2010; Moskowitz & Li, 2011; Moskowitz & Stone, 2012). Instead, goals have the potential to prevent stereotypes from ever becoming activated (Moskowitz, 2010; Moskowitz & Li, 2011; Moskowitz & Stone, 2012). This is important considering that stereotypes often operate unconsciously and automatically; hence, people may not always be aware that stereotype activation has occurred or how stereotype activation might influence subsequent responses (Bargh, Chen & Burrows, 1996; Greenwald & Banaji, 1995; Wilson & Brekke, 1994). Consequently, people may not always utilise reactive stereotype-control strategies.

Empirical evidence supports the influence of goals on stereotype activation (e.g., Macrae, et al., 1997; see Kunda & Spencer, 2003, for a review). Macrae et al. (1997) were some of the first researchers to demonstrate that stereotype activation is contingent on the type of processing goals used. In their first experiment, participants were given one of three processing goals when viewing photographs of female faces and household items. The feature-detection group indicated whether a white dot was present or absent in each photograph. The semantic-judgement group indicated whether the stimulus in each photograph was animate or inanimate. Finally, the exposure group pressed a key each time a new photograph appeared. In between each photograph, participants indicated whether a letter string (either a nonword, or a female stereotypic or counter-stereotypic word) was a word or not. Only participants in the semantic judgement group exhibited activation of the female stereotype (indexed by faster reaction times to female-stereotypic than counter-stereotypic words). Macrae et al. concluded that having a semantic processing goal when processing faces was necessary for stereotype activation to occur.

Subsequently, Galinsky and Moskowitz (2000) demonstrated that stereotype activation is also contingent on a perspective-taking goal. In their first experiment, participants wrote an

essay about a typical day in the life of an elderly man. Before writing the essay, one third of the participants were instructed to take the perspective of the elderly man, one third of the participants were instructed to suppress their stereotypes, and one third of the participants received no additional instructions. Afterward, participants completed a lexical decision task (LDT) to assess accessibility of the elderly stereotype, and wrote a second essay about a typical day in the life of a second elderly man. The results demonstrated that only participants who suppressed stereotypes exhibited accessibility of the elderly stereotype (indexed by faster reaction times to elderly-stereotypic than stereotype-neutral words), suggesting that perspective-taking successfully circumvents stereotype activation. Additionally, participants who suppressed stereotypes, and participants took the perspective of the first elderly man, wrote less stereotypical essays about the second elderly man than participants who received no additional instructions before writing about the first elderly man.

The automatic activation of goals without conscious awareness takes time to establish. According to Bargh and colleagues' auto-motives model (Bargh, 1990; Bargh & Chartrand, 1999), at first, a person may consciously activate a goal in certain situations, but over time, the continued activation of the same goal in the same situations results in a perception-behaviour link. Thus, cues in the environment become linked to the goal, which in turn influences behaviour. As a result, conscious initiation of the goal is no longer needed as cues in the environment are sufficient for triggering the goal automatically. In relation to stereotyping, goals that are inconsistent with stereotyping (e.g., an egalitarian or nonprejudiced goal) may significantly reduce the likelihood of stereotype activation occurring (Kunda & Spencer, 2003; Moskowitz, 2010). For example, people who hold the goal to act egalitarian towards Black people will repeatedly and consistently avoid acting nonegalitarian in the presence of a Black person (Moskowitz et al., 1999; Moskowitz et al., 2000). This goal

may, over time, become linked to all encounters with a Black person, to such an extent that even a perceptual cue like dark skin might come to trigger goal activation rather than stereotype activation (Moskowitz et al., 1999; Moskowitz et al., 2000).

According to Moskwitz and colleagues (Moskowitz et al, 1999; Moskowitz et al., 2000), once conscious awareness is no longer needed for goal activation to occur, the goal becomes chronically accessible. Chronic goals are activated preconsciously and are the dominant response when a person is encountered, preventing the activation of stereotypes altogether. Indeed, Moskowitz and colleagues have demonstrated that chronic egalitarian goals result in less stereotype activation than non-chronic goals (e.g., Moskowitz et al., 1999; Moskowtiz et al., 2000). For example, Moskowitz et al. (1999, Experiment 3) asked participants with a chronic or non-chronic egalitarian goal to pronounce words (female-stereotypic and stereotype-neutral words) following a female or male face. The stimulus onset asynchrony between the presentation of the face and the word was just 200 ms, preventing participants from consciously controlling their responses. Participants with a non-chronic goal pronounced female-stereotypic words faster following stereotype-relevant primes (i.e., female faces) than -irrelevant primes (e.g., male faces), suggestive of stereotype activation. This difference did not occur for participants with a chronic goal. Because conscious control was precluded in the task, Moskowitz et al. concluded that holding a chronic goal prevented stereotype activation from occurring.

In addition to investigating chronic goals, Moskowitz and colleagues (Moskowitz & Li, 2011; Moskowitz & Stone, 2012) have also investigated the effect of temporary egalitarian goals on stereotype activation. For example, Moskowitz and Li (2011, Experiment 1) asked participants to describe a personal failure to act egalitarian towards Black people (temporary egalitarian goal) versus a failure to adhere to traditions. Participants then completed a

reaction-time task to assess stereotype activation. On each trial, participants completed two interspersed tasks. First, participants were presented with a prime (White or Black face) and two letter strings. Participants ignored the face, and indicated whether the letter strings were the same or different. Second, participants were presented with an additional letter string (either a nonword, or a Black-stereotypic or stereotype-neutral word), and indicated whether the letter string was a word or not. If stereotypes are activated, then response times to Black-stereotypic words should be faster following Black faces compared to White faces. Contrary to this, participants with a temporary egalitarian goal demonstrated the exact opposite, namely slower response times to Black-stereotypic words following Black faces compared to White faces. Moskowitz and Li concluded that temporary egalitarian goals can also reduce stereotype activation.

When considering these two findings together, the fact that both chronic and temporary egalitarian goals reduce stereotype activation might seem problematic. Based on Bargh's (1990) auto-motives model, only chronic egalitarian goals should prevent the activation of stereotypes when conscious control is not possible because the very nature of temporarily activated goals implies that the perception-behaviour link has not yet been established, so conscious intent should still be necessary for goal activation. However, Moskowitz et al. (2000) argued that goals to which people are committed (whether chronically accessible or temporarily activated) result in stereotype control; hence, goal commitment, rather than the chronicity of the goal, may be the key determinant of whether stereotype activation occurs.

Although it is clear from past research that goals influence stereotype activation (see Kunda & Spencer, 2003, for a review), less attention has focused on the mechanism behind that influence. This suggests that a refinement of our understanding about how goals influence

stereotype activation is needed, as well as the development of second-generation theories investigating the processes and mechanisms. In the next two sections of this chapter, I will review evidence that provides insight into the mechanism behind the influence of goals on stereotype activation and prejudice, including the role of goal source and goal–behaviour discrepancies.

1.5 Goal Source and Stereotype Control

Plant and Devine (1998) took a significant step towards identifying the mechanism behind the influence of goals on stereotype use. They distinguished between two types of self-regulatory prejudice-reduction goals—namely, internal motivation to respond without prejudice (arising from internalised and personally endorsed nonprejudiced standards) and external motivation to respond without prejudice (arising from social pressure to conform to nonprejudiced standards, such as the politically correct (PC) standards). They argued that a person could be primarily motivated by internal reasons, external reasons, a mixture of internal and external reasons, or not particularly motivated.

To assess this distinction, Plant and Devine (1998, Experiments 1 & 2) developed two five-item scales assessing internal and external motivation. The internal motivation scale (IMS) assessed whether participants' motivation to act nonprejudiced is based on their own nonprejudiced beliefs and values (e.g., "I am personally motivated by my beliefs to be nonprejudiced toward Black people"). In contrast, the external motivation scale (EMS) assessed whether participants' motivation to act nonprejudiced is based on social pressure (e.g., "I try to act nonprejudiced toward Black people because of pressure from others"). Analysis of participants' responses to the IMS and EMS revealed good validity of the scales. Plant and Devine found that the IMS was correlated strongly with explicit measures of prejudice, indicating that internal motivation is associated with lower prejudice. The EMS

was also correlated with explicit measures of prejudice, but only weakly, and in the opposite direction, indicating that external motivation was somewhat associated with greater prejudice.

Plant and Devine (1998, Experiment 3) also provided evidence for the validity of the IMS and EMS. They found that participants high in internal and low in external motivation (chronic internal motivation) exhibited low endorsement of Black stereotypes, regardless of whether they responded anonymously in a private setting or publicly in front of the experimenter. However, participants low in internal and high in external motivation (chronic external motivation) exhibited lower endorsement of Black stereotypes when responding in public than in private. This finding demonstrated that externally motivated participants conform to the nonprejudiced norm when they are concerned about how others perceive them. Taken together, Plant and Devine's experiments demonstrate that goal source plays an important role in determining whether stereotyping and prejudice occurs.

Plant and Devine (1998) are not the only researchers who have attempted to explicate the mechanism behind the influence of goals on prejudice and stereotyping. For example, Fazio et al. (1995) argued that whether or not an individual responds in a nonprejudiced manner towards an outgroup member is determined by the valence of the automatically activated of evaluation and the extent to which the individual attempts to control the automatic evaluation. Nonprejudiced individuals experience no automatic activation of negative evaluations in response to an outgroup member because they have replaced the automatic negative beliefs with their own personal beliefs (e.g., automatic positive evaluations). In contrast, prejudiced individuals experience automatic activation of negative evaluations in response to encountering an outgroup member, and because their personal beliefs are consistent with the automatic negative evaluations, they do not attempt to control the automatic negative evaluations. Finally, the remaining individuals are characterised by a

conflict between the automatic activation of negative evaluations in response to encountering an outgroup member and their personal beliefs, which are inconsistent with the automatic negative evaluations (i.e., they are nonprejudiced). Fazio et al. (1995) argued that these individuals are concerned with acting nonprejudiced either because of a personal discomfort with the automatic negative evaluation or a desire to avoid appearing prejudiced to others. The key difference between Plant and Devine's (1998) and Fazio et al's (1995) accounts is that Fazio et al. do not distinguish between whether the latter group of individuals act nonprejudiced for internal or external reasons, but rather argue that these individuals are generally concerned with acting nonprejudiced.

Lastly, the work in the present thesis examines how goal source and goal–behaviour discrepancies influence stereotype activation. Although it is intuitively reasonable to expect that the self-regulation of stereotyping would be directed toward controlling stereotype use—after all, the purported goal is either to behave in a more egalitarian manner (for internally motivated individuals) or to be seen to do so (for externally motivated individuals)—the impact of self-regulatory motives on stereotype activation may be less clear. Nonetheless, there is reason to expect that prejudice-related self-regulatory goals will influence stereotype activation. Plant and Devine (2009) have demonstrated that people who are internally motivated to act nonprejudiced intend to eradicate all forms of prejudice. These individuals seek to eradicate not only prejudiced behaviour, but also prejudiced thoughts and feelings—that is, emotional reactions and stereotypes. This suggests that to the extent that individuals seek to eradicate stereotypes (whether because they personally want to or because they are expected to do so by others), they will be vigilant toward not only goal–behaviour discrepancies, but also goal–thought discrepancies.

1.5.1 Internal motivation. Empirical research has demonstrated that people high in internal motivation exhibit low levels of explicit and implicit prejudice towards Black people (e.g., Amodio, Harmon-Jones, & Devine, 2003; Devine et al., 2002; Hausmann & Ryan, 2004; Legault, Gutsell, & Inzlicht, 2011). This reduction in prejudice for people high in internal motivation is not specific to Black targets, however. Research has shown that men high in internal motivation to respond without sexism towards women (e.g., Klonis, Plant, & Devine, 2005), and heterosexuals high in internal motivation to respond without prejudice towards homosexuals and lesbians (e.g., Ratcliff, Lassiter, Markman, & Snyder, 2006), both exhibit low levels of prejudice. In sum, research has repeatedly established that high levels of internal motivation reduce prejudice (see Butz & Plant, 2004, for a review).

High internal motivation is also associated with lower stereotype endorsement (e.g., Plant & Devine, 1998; Plant, Devine, & Brazy, 2003) and less stereotype activation (e.g., Johns, Cullum, Smith, & Freng, 2008; Peruche & Plant, 2006). For example, Johns et al. (2008, Experiment 2) found that as internal motivation increased relative to external motivation, stereotype activation decreased (i.e., less facilitation to Black-stereotypic compared to stereotype-neutral words following a Black face on an LDT). Further, mediational analysis demonstrated that the reduction in stereotype activation for participants higher in internal motivation was significantly mediated by the automatic activation of an egalitarian goal (see also Moskowitz, 2002; Moskowitz et al., 2011; Moskowitz & Stone, 2012).

Amodio, Devine, and Harmon-Jones (2008, Experiment 1) have argued that the low levels of stereotype activation and prejudice exhibited by people high in internal motivation on implicit tasks may be due to greater self-regulatory control. In their research, participants with varying degrees of internal and external motivation indicated whether an object was a

gun or a tool, following primes of White and Black faces (adapted weapons identification task; Payne, 2001). Process dissociation analysis demonstrated that participants high in internal and low in external motivation exhibited greater self-regulatory control over their responses compared to participants high in both internal and external motivation. Amodio et al. also investigated whether the greater self-regulatory control is due to greater conflict-monitoring (i.e., monitoring conflict between personal beliefs and unwanted stereotypes). Analysis revealed that participants high in internal and low in external motivation exhibited significantly greater error-related negativity (ERN; a component of an event-related electrocortical response and an accepted indicator of conflict monitoring; see Amodio et al., 2008) when stereotype inhibition was required (i.e., on Black–tool, but not Black–gun combinations) compared to participants high in both internal and external motivation. Furthermore, greater conflict monitoring on Black–tool trials explained the difference in self-regulatory control between participants high in internal and low in external motivation, and participants high in both internal and external motivation. In sum, high internally motivated people are able to respond with low levels of stereotype activation and prejudice on implicit tasks because of greater self-regulatory control due to monitoring for unwanted stereotypes.

The greater self-regulatory control that people high in internal motivation demonstrate also increases the effectiveness of reactive stereotype-control strategies. For example, research has shown that high internal motivation allows participants to circumvent the stereotype rebound effect (Gordijn et al., 2004; Wyer, 2007). Gordijn et al. (2004, Experiment 3) demonstrated that although participants low in internal motivation to suppress stereotypes exhibited the stereotype rebound effect (indexed by faster reaction times to skinhead-stereotypic compared to stereotype-neutral words after successfully suppressing the skinhead stereotype while writing about a typical day in the life of a skinhead), participants high in

internal motivation to suppress stereotypes did not exhibit the stereotype-rebound effect. Furthermore, Wyer (2007, Experiment 2) demonstrated that only participants high in external and low in internal motivation exhibited the stereotype-rebound effect compared to all other internal/external motivation combinations. People with high internal motivation circumvent the stereotype-rebound effect ostensibly because their egalitarian goal is automatically activated by cues in the environment that signal stereotyping or prejudice might occur, just like chronic goals (Hausmann & Ryan, 2004; Monteith, Sherman, & Devine, 1998; Moskowitz et al., 1999; Moskowitz et al., 2000; see also Bargh's (1990) auto-motives model). As a result, stereotype control is less taxing of cognitive resources for these individuals; hence, stereotypes do not rebound while cognitive resources are being restored (Gordijn et al., 2004; Hausmann & Ryan, 2004).

1.5.2 External motivation. Less empirical research has investigated how external goals influence stereotype activation and prejudice. In addition, empirical investigations examining external goals find inconsistent results. Some empirical research, particularly where external pressure is induced via the presence of an audience, has shown that participants comply with the nonprejudiced goal (i.e., exhibit nonprejudiced behaviour; e.g., Blanchard, Lilly, & Vaughn, 1991; Castelli & Tomelleri, 2008; Monteith et al., 1996; Plant & Devine, 1998, 2001; Plant et al., 2003). For example, Castelli and Tomelleri (2008, Experiment 1) asked participants to complete an IAT (Greenwald et al., 1998) alone or in a group of three. Analysis revealed that the participants in the public condition (in a group of three) had lower IAT scores, indicating less prejudice (the higher the IAT scores the higher the association between Black and unpleasant compared to White and unpleasant). Castelli and Tomelleri argued that the public condition automatically activated egalitarian social

norms whereas the private condition did not. They concluded that prejudiced responses can be controlled by participants when in the presence of others.

Although the above research suggests that participants may comply with external pressure to act nonprejudiced at times, this is not consequence-free. Plant and Devine (2001) investigated how people react to external pressure to act nonprejudiced. In Experiment 3, Plant and Devine recruited participants who had opposed a new recruitment strategy designed to attract Black students using scholarships. These participants were informed that the committee was recruiting essays for and against the policy, but had received too many essays opposing the recruitment strategy; hence, participants were requested to write a counterattitudinal essay favouring the recruitment strategy. This methodology activates two competing norms: a descriptive norm against the recruitment strategy because most students wrote essays opposing the recruitment strategy, and an injunctive norm favouring the recruitment strategy because society demands that people act in a pro-Black manner (or, at least, in an egalitarian manner). Only 9% of 111 participants did not comply with the external request to write an essay favouring the recruitment strategy, which was not affected by participants' motivation (i.e., internal or external motivation). For the remaining participants, after complying with the external request, participants low in internal and high in external motivation exhibited more attitudinal and behavioural backlash (i.e., greater opposition to the policy) than all other internal/external motivation combinations. Furthermore, the increased angry and threatened affect that participants low in internal and high in external motivation experienced as a result of writing the essay partially mediated the increased backlash. Therefore, just like stereotype suppression and the subsequent stereotype rebound effect, externally motivated participants may comply with external pressure to act nonprejudiced at first, but subsequently exhibit prejudiced behaviour.

Additional empirical research has also shown that externally motivated participants may fail to comply with pressure to act nonprejudiced (i.e., exhibit prejudiced behaviour; e.g., Devine et al., 2002; Hausmann & Ryan, 2004; Legault et al., 2011; Peruche & Plant, 2006; Plant & Devine, 2001). For example, Legault et al. (2011, Experiment 2) primed participants with either an autonomous motivation or a controlled motivation. Specifically, one group of the participants were primed with an autonomous motivation (i.e., internal; e.g., "Being nonprejudiced is important to me") to act in a nonprejudiced manner before writing three sentences about why acting nonprejudiced is personally satisfying or important. A second group of participants were primed with a controlled motivation (i.e., external; e.g., "I should avoid being a racist") to act in a nonprejudiced manner before writing three sentences about the social expectation or obligation to act nonprejudiced. A control group were not primed with any motivation. Participants then completed an explicit measure of prejudice (Symbolic Racism Scale; Henry & Sears, 2002) and an implicit measure of prejudice (IAT). The results showed that participants primed with internal (autonomous) motivation to act nonprejudiced exhibited less explicit and implicit prejudice towards Black people than participants not primed with motivation. In contrast, participants primed with external (controlled) motivation to act nonprejudiced exhibited greater explicit and implicit prejudice towards Black people than participants not primed with motivation; hence, participants primed with external motivation to act nonprejudiced did not comply with pressure to act nonprejudiced on the explicit or implicit prejudice measures.

There is no clear-cut explanation for why externally motivated people sometimes fail to comply with an external goal to act nonprejudiced. One argument suggests that externally motivated people resent the restriction placed on their freedom and so act inconsistently with the external goal (i.e., exhibit prejudiced behaviour) as a means of rebelling against the people

responsible for imposing the attitude (see Brehm, 1966; Legault et al., 2011; Plant & Devine, 2001). Reactance does not occur for internally motivated people because the external pressure is consistent with their own personal beliefs and values about acting nonprejudiced (Plant & Devine, 2001). Indeed, Plant and Devine (2001, Experiment 1) found that participants low in internal and high in external motivation felt more constrained by and more resentment towards external pressure than any other internal/external motivation combination.

An alternative argument, proposed by Hausmann and Ryan (2004), is that when externally motivated participants are faced with external pressure to act egalitarian, they may be motivated to control their automatic prejudices but unable to do so. Controlling automatic prejudice on implicit prejudice measures like the IAT might be cognitively demanding for externally motivated people because they do not have as much experience controlling their prejudice as internally motivated people and controlling responses on reaction time measures is difficult (Hausmann & Ryan, 2004). As a result, externally motivated people's cognitive resources may become depleted and their attempt at controlling their automatic prejudice may backfire (i.e., increased prejudicial responses rather than decreased). In support of this argument, Hausmann and Ryan found that external motivation was positively associated with implicit prejudice (assessed through the IAT); in other words, as external motivation increased, implicit prejudice towards Black people also increased. In contrast, internal motivation was negatively associated with implicit prejudice (assessed through the IAT); in other words, as internal motivation increased, implicit prejudice towards Black people decreased. In addition, externally motivated participants reported effortful control during the IAT and this alone mediated the relationship between external motivation and increased prejudice; neither general concern about appearing prejudiced nor anxiety while completing the task mediated the relationship.

Either of these accounts could explain why externally motivated people sometimes do not comply with external pressure to act egalitarian. This thesis aimed to further our understanding of the mechanism through which external goals influence stereotype activation. External goals and internal goals are equally important considering that people may begin with an external goal that gradually becomes internalised (see self-determination theory; Deci & Ryan, 1985; see also Ryan & Deci, 2000). Understanding how external goals function may lead to strategies that aid the conversion of an external goal to an internal goal.

1.6 Goal Source, Goal–Behaviour Discrepancies, and Stereotype Control

In addition to goal source, the size of the discrepancy between people's actual behaviour and their standards for behaviour may also be an important component of the mechanism through which goals influence stereotype activation and prejudice. Inspired by self-discrepancy theory (Higgins, 1987), Plant and Devine (1998) argued that discrepancy size (distance between actual behaviour and desired standards) interacts with internal and external motivation to influence affect. In Experiment 3, participants read five scenarios about how people should think or feel in situations involving Black people. For example, in one scenario, participants read, "Imagine that you saw a young Black woman at the grocery store with four small children. Your initial thought should be—'How typical'," (Plant & Devine, 1998, p. 819). For each scenario, participants indicated how they personally believed they *should* respond (personal standards) or how they *should* respond according to campus standards (other-imposed standards), and how they *would* respond (predicted actual behaviour). In addition, participants completed an affect measure assessing their reactions to violating personal or other-imposed standards. The findings revealed that for people high in internal motivation, larger (vs. smaller) discrepancies between actual responses and personal standards resulted in higher levels of guilt. In contrast, for people high in external motivation, larger (vs.

smaller) discrepancies between actual responses and other-imposed standards resulted in higher levels of threatened affect. This experiment demonstrates that the nature of the goal–behaviour discrepancy has important consequences for affect depending on whether a person is internally or externally motivated.

Plant and Devine's (1998) evidence was restricted to affective responses. They speculated, but did not test, how goal–behaviour discrepancies in the context of acting nonprejudiced might influence self-regulation when a person is internally or externally motivated. For people who are internally motivated, Plant and Devine suggested that a large goal–behaviour discrepancy might prompt goal-consistent behaviour (i.e., reduced prejudicial responses) designed to reduce the discrepancy. Because internally motivated people value acting nonprejudiced and freely adopt the nonprejudiced attitude, this attitude is inherently strong and central (Sherman & Gorkin, 1980). The nonprejudiced attitude may subsequently become internalised (i.e., high commitment and importance) and ultimately self-defining (i.e., part of the self-concept; Devine, Monteith, Zuwerink, & Elliot, 1991; Monteith, Devine, & Zuwerink, 1993; Plant & Devine, 1998; Sherman & Gorkin, 1980). Failing to adhere to their nonprejudiced attitude (actual failure or the potential for future failure) may threaten the self-concept of these individuals (Devine et al., 1991; Dutton & Lake, 1973; Monteith, 1993; Monteith et al., 1993; Sherman & Gorkin, 1980; Wicklund & Gollwitzer, 1982). This threat motivates these individuals to reduce the goal–behaviour discrepancy and bolster their nonprejudiced attitude (Dutton & Lake, 1973; Sherman & Gorkin, 1980; Wicklund & Gollwitzer, 1982), which can be achieved by acting consistently with the nonprejudiced attitude (Monteith, 1993; Wicklund & Gollwitzer, 1982).

For people who are externally motivated, however, Plant and Devine (1998) speculated that a large goal–behaviour discrepancy in the context of acting nonprejudiced

might result in a lack of goal-consistent or even goal-inconsistent behaviour (i.e., prejudicial or increased prejudicial responses). Because externally motivated people do not freely adopt the nonprejudiced attitude, the attitude is not as strong or central (Sherman & Gorkin, 1980), and is less internalized (i.e., commitment and importance are lower), meaning that the nonprejudiced attitude is not self-defining (i.e., part of the self-concept; Devine et al., 1991; Monteith et al., 1993; Plant & Devine, 1998). Failing to adhere to their nonprejudiced attitude (actual failure or the potential for future failure) has no implications for the self (Devine et al., 1991; Monteith, 1993; Monteith et al., 1993). Rather than attempting to bolster their nonprejudiced attitude, these individuals may reject the attitude entirely and perhaps even act consistently with the goal–behaviour discrepancy (i.e., become more prejudiced).

Plant and Devine's (1998) speculation suggests that both the source of the goal (i.e., internal vs. external) and goal–behaviour discrepancies are important. However, it is unclear from Plant and Devine's speculation whether the magnitude of the goal–behaviour discrepancy matters for stereotype activation and prejudice. Plant and Devine's empirical research demonstrates that small and large goal–behaviour discrepancies have different consequences for *affect* depending on goal source, but they only speculate about how large goal–behaviour discrepancies will influence stereotyping and prejudice. It is unclear, therefore, whether small goal–behaviour discrepancies will affect stereotype activation and prejudice differently depending on goal source.

While research examining the mechanism through which egalitarian goals influence stereotype activation and prejudice is lacking, the self-regulation literature has investigated how goals influence self-regulation in greater detail. This theory and empirical evidence have not been applied to stereotype activation. Therefore, in an effort to explore how goals

influence stereotype activation, I aimed to apply theory and evidence from the self-regulation literature to stereotype activation.

1.7 Self-Regulation Models

Fishbach and colleagues (Fishbach & Dhar, 2005, 2007; Fishbach et al., 2009; Koo & Fishbach, 2008; see Figure 1) have proposed a model of goal progress and motivation. The model asserts that motivation and goal adherence are a function of the goal–behaviour discrepancy between actual and desired goal states, and whether movement towards a goal is construed in terms of progress or commitment to the goal. Here, I adopt a basic version of the model that focuses on whether goals are internally generated (e.g., derived from personal beliefs and values) or externally imposed (e.g., by the government). For a more thorough treatment of the model, see Appendix A.

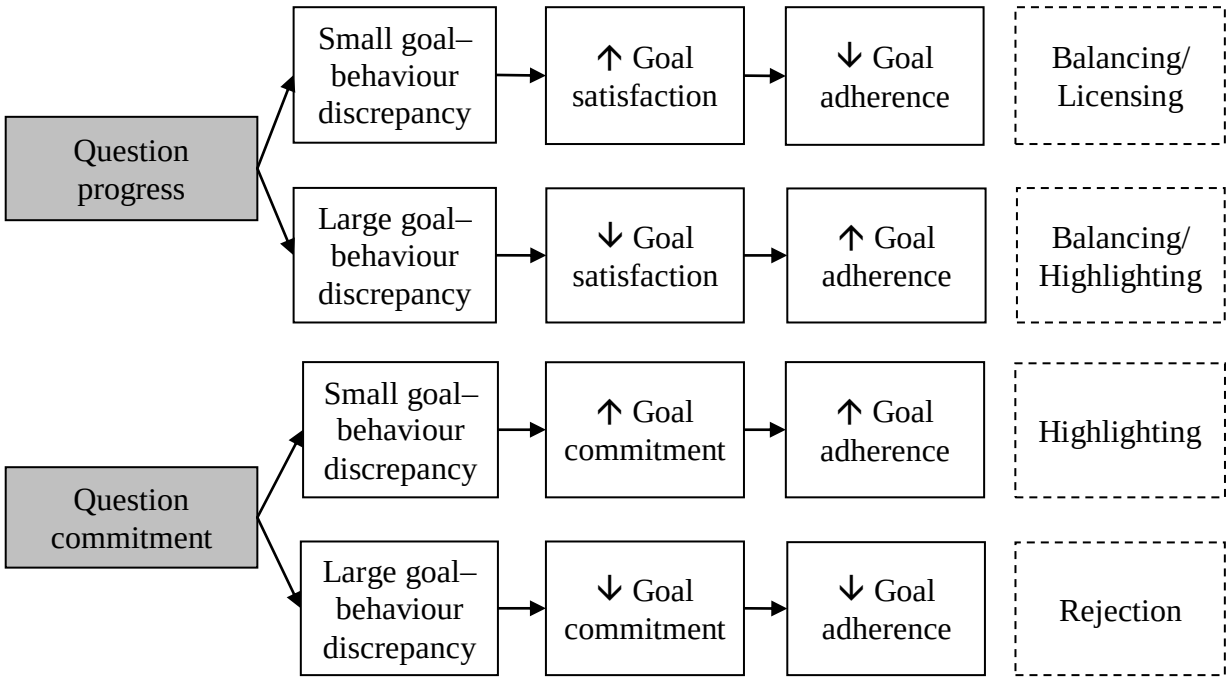

***Figure 1*. Diagram of a simplified version of Fishbach and colleagues' model of self-regulation.**

For internally generated goals, commitment is usually unambiguously certain, leading people to ask questions about their goal progress (asking questions about commitment would provide redundant information; Fishbach et al., 2009; Koo & Fishbach, 2008). When people represent a goal action as signalling progress, they ask "Am I making sufficient progress towards this goal?" and assess how satisfied they are with the discrepancy between the current goal state and the desirable state (Fishbach & Dhar, 2005; Fishbach, et al., 2009). When people ask about progress, they tend to follow a dynamic of *balancing*, whereby a goal-congruent action subsequently "licenses" the individual to disengage and pursue other, even contradictory, goals (Fishbach & Dhar, 2005, 2007). Therefore, a small goal–behaviour discrepancy signals that sufficient progress has been made towards the desired end-state, resulting in decreased goal adherence and attention turning to other, even contradictory, goals that have been neglected. In contrast, a large goal–behaviour discrepancy signals that insufficient progress has been made towards the desired end-state, resulting in increased goal adherence, promoting subsequent goal-consistent actions.

For externally generated goals, commitment is usually uncertain, leading people to ask questions about their commitment towards the goal (Fishbach et al., 2009; Koo & Fishbach, 2008). When people represent a goal action as signalling commitment, they ask "Am I committed to this goal?"—using self-perception of their own behaviour—and assess how valuable the goal is and the likelihood of goal attainment (Fishbach et al., 2009; Koo & Fisbach, 2008). As a result, people tend to follow a dynamic of *highlighting*, whereby a goal is prioritised for completion (Fishbach & Dhar, 2005, 2007). When people ask about commitment, a small goal–behaviour discrepancy signals high commitment to the goal, resulting in increased goal adherence, promoting subsequent goal-consistent actions. In contrast, a large goal–behaviour discrepancy signals low commitment to the goal, resulting in

decreased goal adherence (i.e., lack of goal-consistent action in this case as the external goal is rejected).

There are alternative models of self-regulation that might potentially explain the mechanism through which goals influence stereotype activation and prejudice. For example, regulation focus theory (Higgins 1997, 1998) proposes that people can adopt either a promotion-focused strategy, which emphasises approaching positive outcomes (e.g., behaviour matches the desired end-state of a goal), or a prevention-focused strategy, which emphasises avoiding negative outcomes (e.g., behaviour mismatches with the desired end-state of a goal). Additionally, the goal-shielding model (Kruglanski et al., 2002; Shah, Friedman, & Kruglanski, 2002; Shah & Kruglanski, 2002) proposes that when commitment to a goal is high, goal pursuit is aided by "facilitating the detection and processing of goal-relevant stimuli (as well as other goals)" (Moskowitz & Li, 2011, p. 3) and inhibiting incompatible goals. Finally, an implementation intention strategy (Gollwitzer, 1999) proposes that forming an "if-then" plan, which links a situation with a goal-directed response, improves goal attainment. Although regulatory focus theory highlights the importance of goal–behaviour discrepancies (Higgins, 1987), and may be applicable to the internal/external goal source distinction in that previous research has linked "ideal" standards to one's own goals and "ought" standards to the goals set by others (Plant and Devine, 2009), it does not speak as clear to how these different goal sources interact with discrepancies to prompt control versus licensing. To the best of my knowledge, there is no research linking the goal-shielding model or an implementation intention strategy to goal source. While the goal-shielding model and an implementation intention strategy could be adapted to take goal source into account, it is nonetheless the case that the role of goal source and discrepancy size is already integrated into Fishbach and colleagues' model. Given that the work of Plant and Devine (1998) and

Moskowitz and colleagues (Moskowitz & Li, 2011; Moskowitz & Stone, 2012) focus heavily on goal source, and this work provides the basis for the work in the present thesis, Fishbach and colleagues' model is the most easily amenable.

2.0 The Present Research

The majority of empirical investigations examining internal goals and goal–behaviour discrepancy size have focused on prejudice rather than stereotype activation (except see Moskowitz & Li, 2011; Moskowitz & Stone, 2012). Additionally, there has been no systematic investigation of the potential influence of external goals and goal–behaviour discrepancy size on stereotype activation. Only one investigation has examined the effect of external motivation on prejudice following a self-regulatory failure. Fehr and Sassenberg (2010, Experiment 2) found no significant interaction between internal motivation, external motivation, and self-regulatory failure. Instead, only those motivated internally to act nonprejudiced responded with less prejudice towards Arabs following a self-regulatory failure versus no-failure; this finding supports Plant and Devine's (1998) speculation about internally motivated individuals only. As a result, the mechanism through which internally-generated and externally-imposed goals influence stereotype activation and prejudice remains unclear.

In an attempt to further understand this mechanism, this thesis aimed to investigate the influence of both internal and external goals, and small and large goal–behaviour discrepancies, on stereotype activation. I applied Fishbach and colleagues' model to stereotype activation. For an internal egalitarian goal, perceiving a large goal–behaviour discrepancy should result in increased goal adherence, leading people to exhibit control over stereotype activation, whereas perceiving a small goal–behaviour discrepancy should result in decreased goal adherence, leading people to exhibit stereotype activation. For an external egalitarian goal, however, perceiving a large goal–behaviour discrepancy should result in

decreased goal adherence, leading people to exhibit stereotype activation, whereas perceiving a small goal–behaviour discrepancy should result in increased goal adherence, leading people to exhibit control over stereotype activation.

It is important to consider whether decreased adherence to an egalitarian goal and the subsequently predicted stereotype activation represents a lack of goal-consistent action or goal-inconsistent action. People with external goals follow a dynamic of highlighting where a single goal is prioritised for completion above all other goals (Koo & Fishbach, 2008); therefore, when a large goal–behaviour discrepancy signals that goal commitment is low, goal adherence decreases because the external goal is rejected. Consequently, these individuals simply do not respond consistently with the rejected external goal (i.e., stereotype activation signals a lack of goal-consistent action). People with internal goals, however, following a dynamic of balancing amongst multiple goals (Koo & Fishbach, 2008); therefore, when a small goal–behaviour discrepancy signals that sufficient goal progress has been made towards the goal, this licenses the disengagement from current goal pursuit and the pursuit of other goals. When these individuals subsequently pursue a goal that is unrelated to egalitarianism, effort to control stereotypes ceases and stereotype accessibility may simply be restored to its original level of activation prior to goal pursuit (i.e., stereotype activation signals a lack of goal-consistent action). However, it is feasible that these individuals may subsequently pursue a goal that is contradictory to egalitarianism. Positive distinctiveness may serve as a competing goal that internally motivated people pursue following a small goal–behaviour discrepancy. According to social identity theory (Tajfel & Turner, 1979, 1986), people are motivated to form a positive social identity for their ingroup. This entails making positive downward comparisons to relevant outgroups so that one's ingroup is positively distinct from relevant outgroups (Tajfel & Turner, 1979, 1986). Successful egalitarian goal pursuit may

highlight the similarities between the ingroup and the outgroup for internally motivated people, which may exceed these individuals desire for assimilation with the outgroup (see optimal distinctiveness theory; Brewer, 1991). As a result, these individuals may be motivated to re-establish their uniqueness and restore the positive distinctiveness of their ingroup relative to the outgroup (Brewer, 1991).

Applying Fishbach and colleagues' model may increase our understanding of the mechanism through which internally-generated and externally-imposed goals influence stereotype activation. However, it is important to briefly discuss whether Fishbach and colleagues' model is appropriate within an egalitarian context. On the one hand, there is reason to believe that Fishbach and colleagues' model is appropriate within an egalitarian context. Comparing Plant and Devine's theory (1998) to Fishbach and colleagues' model reveals that the theoretical positions are congruent with each other for large goal–behaviour discrepancies. They both argue that people with an internal goal and a large goal–behaviour discrepancy will exhibit goal-consistent behaviour, whereas people with an external goal and a large goal–behaviour discrepancy will exhibit a lack of goal-consistent or even goal-inconsistent behaviour. Admittedly, this is based on speculation from Plant and Devine rather than empirical evidence. In addition, Fishbach and colleagues' model suggests that small goal–behaviour discrepancies may differentially influence goal behaviour; namely, people with an internal goal and a small goal–behaviour discrepancy should exhibit a lack of goal-consistent or even goal-inconsistent behaviour, while people with an external goal and a small goal–behaviour discrepancy should exhibit goal-consistent behaviour. The differential effect of small goal–behaviour discrepancies on stereotype activation depending on internal and external motivation has not yet been investigated within the stereotyping and prejudice literature. Applying Fishbach and colleagues' model to stereotype activation may provide

some useful insight into how internally-generated and externally-imposed goals influence stereotype activation.

On the other hand, there are two caveats which suggest that Fishbach and colleagues' model may not be appropriate within an egalitarian context. First, Fishbach and colleagues' evidence is based on research examining goals with clear-end states where an individual can say with absolute certainty that a goal has been achieved (e.g., a diet goal to lose 10 lbs is achieved when a person has lost 10 lbs). Egalitarian goals, in contrast, do not necessarily have such clear-end states because these goals are continuously pursued, with each new encounter with an out-group member offering a new opportunity to act consistently or inconsistently with the egalitarian goal. If all egalitarian goals do lack clear-end states then goal–behaviour discrepancy size is unlikely to have an effect on the self-regulation of egalitarianism (Fishbach, 2009).

The second caveat that should be noted regarding the appropriateness of Fishbach and colleagues' model within an egalitarian context relates to the model's implicit assumption that internal and external motivation are mutually exclusive. Yet, Plant and Devine (1998) demonstrated that people can be motivated to act nonprejudiced primarily for internal reasons (i.e., have an internal nonprejudiced goal), primarily for external reasons (i.e., have an external nonprejudiced goal), or for both internal and external reasons (i.e., have both an internal and an external nonprejudiced goal). An interesting question arises then: If a person holds both an internal and an external egalitarian goal, what determines whether they will focus on their progress towards the goal, or their commitment to the goal?

According to Plant and Devine (1998, 2009), when people are both internally and externally motivated, internal motivation is the primary motivation. In line with this conclusion, Plant and Devine (2009) demonstrated that while individuals high in both internal

and external motivation had both the intention to eradicate prejudice altogether (also held by individuals high in internal but low in external motivation) and the intention to hide prejudice from others (also held by individuals low in internal but high in external motivation), the intention to eradicate prejudice altogether was primary. When given the opportunity to use a program that would reduce detectable prejudice in the short-term but inadvertently increase undetectable prejudice in the long-term, individuals high in both internal and external motivation were not interested in the program. The research by Plant and Devine (2009) suggests that if a person holds an internal and an external goal we should expect them to ask questions about their progress towards the goal and demonstrate the same self-regulatory pattern for internal goals within Fishbach and colleagues' model. This empirical question was not one of the primary questions addressed within the present thesis, but is an important empirical question for future research.

Finally, this thesis also aimed to determine how early egalitarian goals take effect. Although much research has examined whether an egalitarian influences stereotype activation and use, there has been little consideration of whether an egalitarian goal also influences social categorisation. Yet, in order to determine *how* goals influence stereotype activation and control, it is important to establish *when* goals take effect. On the one hand, egalitarian goals may not influence social categorisation at all. Egalitarian goals may only be activated when there is some implicit or explicit recognition that stereotyping might occur (e.g., Wilson & Brekke, 1994). On the other hand, egalitarian goals may affect whether targets at this stage are categorised as members of the stereotyped group.

The empirical work is divided into three chapters. Chapter 2 examined the effect of priming internal and external egalitarian goals on stereotype accessibility after participants reflecting on a past successful act of egalitarianism. Chapter 3 examined the effect of priming

internal and external goals on stereotype accessibility after participants received false feedback that highlighted the potential for a large or small discrepancy between their actual behaviour and the future likelihood of acting egalitarian. Chapters 2 and 3 also examined the potential moderating role of identification and goal importance. Finally, Chapter 4 extended Chapters 2 and 3 by investigating whether goals effect social categorisation. The effect of goal source and goal–behaviour discrepancy size on social categorisation has not been investigated. Fishbach and colleagues' model may offer useful insights into how goals influence social categorisation.

CHAPTER 2

CONTEMPLATING SUCCESS AND EXTERNAL EGALITARIAN GOALS DECREASE STEREOTYPE ACCESSIBILITY

In this chapter, I examined the effect of reflecting on a past successful egalitarian act on stereotype activation. Research has predominantly examined the effect of failed egalitarian goal pursuit on stereotype activation and prejudice. However, examining how successful egalitarian goal pursuit influences stereotype activation and prejudice may provide a more complete understanding of when people are more or less likely to act egalitarian. Yet, to the best of my knowledge, there has been no empirical examination of how successful egalitarian goal pursuit influences stereotype accessibility for people with external egalitarian goals. In Experiment 1, therefore, I examined how the source of an egalitarian goal (i.e., internal vs. external) influenced stereotype accessibility among individuals who described a past successful egalitarian act. Specifically, half of the participants were primed with internal reasons (personal beliefs and values) to act in an egalitarian manner before reflecting on a time when they successfully acted egalitarian towards a Black person because they felt personally compelled to do so (internal goal source). The other half of the participants were primed with external reasons (other people's expectations) to act in an egalitarian manner before reflecting on a time where they successfully acted egalitarian towards a Black person because they felt socially/politically influenced to do so (external goal source). Results from an LDT indicated that stereotypes were subsequently accessible for participants in the internal, but not the external, goal condition. This finding suggests that participants in the external goal condition had better stereotype control, and is consistent with the interpretation that participants in the internal goal condition relaxed their goal pursuit in the face of evidence

of goal progress. This finding has important implications for our understanding of the impact of success on self-regulation for internal and external egalitarian goals.

1.0 General Introduction

The work in this chapter aimed to further our understanding about how successful egalitarian goal pursuit influences stereotype activation in order to provide a more holistic understanding of when people are more or less likely to act egalitarian.

1.1 Ironic Effect of Success on Prejudice Regulation

Evidence within the prejudice literature indicates that establishing oneself as egalitarian—by successfully acting egalitarian or by considering the prejudiced acts that one could have, but did not, commit—has important implications for subsequent prejudice regulation (e.g., Bradley-Geist, King, Skorinko, Hebl, & McKenna, 2010; Effron, Cameron, & Monin, 2009; Effron, Miller, & Monin, 2012; Monin & Miller, 2001). For example, Monin and Miller (2001, Experiment 2) found that after choosing to hire a Black man (vs. a White man) for a position in a consulting firm, participants were significantly more likely to favour choosing a White man (vs. Black man) for a job as an officer in a small town. Monin and Miller argued that establishing oneself as nonprejudiced (earning "moral credentials") reduces the fear of being labelled a racist and licenses the expression of a pro-White (anti-Black) attitude that is inconsistent with being nonprejudiced.

More recently, Effron and colleagues (2009) have argued that establishing oneself as egalitarian does not license the expression of prejudice per se, but rather builds an expectation that subsequent actions favouring White people in ambiguous situations will not reflect, or at least not be interpreted as, prejudice. Indeed, because the officer hiring decision in Monin and Miller's (2001) task was characterised by racial tension, participants may have chosen a

White applicant for the officer's position to avoid subjecting the Black applicant to hostile working conditions (Effron et al., 2009). Therefore, establishing oneself as egalitarian (earning moral credentials) may allow people to act in a pro-White (anti-Black) manner in ambiguous situations where extenuating circumstances make a pro-White preference more acceptable (Effron et al., 2009). No empirical investigation has examined whether establishing oneself as egalitarian results in the expression of constructs other than prejudice (e.g., stereotype activation), however.

1.2 Ironic Effect of Success for Internal Goals

Fishbach and colleagues' model (Fishbach & Dhar, 2005, 2007; Fishbach et al., 2009; Koo & Fishbach, 2008; see Figure 1) proposes that perceiving a small goal–behaviour discrepancy—which may be more likely following successful goal-directed behaviour—has a different impact on self-regulation for people with internal versus external goals. Fishbach and colleagues propose that because commitment to an internal goal is certain, individuals with internal goals question their progress. Perceiving a small goal–behaviour discrepancy signals that sufficient progress has been made towards the goal, "licensing" these individuals to disengage from the current goal pursuit in order to pursue other, even contradictory, goals. As a result, perceiving a small goal–behaviour discrepancy in relation to an internal egalitarian goal may result in stereotype activation.

Only one program of research by Moskowitz and Li (2011; see also Moskowitz & Stone, 2012) has examined the impact of reflecting on past successful egalitarian behaviour (one source of information about goal–behaviour discrepancies) on stereotype activation for people with an internal goal. For example, in Experiment 3, participants rated the importance of an egalitarian goal (among 25 other goals and values) and subsequently reflected on a time where they successfully acted egalitarian or failed to act egalitarian towards a Black person.

Participants then completed a primed LDT; on each trial, participants memorised White or Black faces (primes) and then indicated whether a letter string (either a nonword, or a Black-stereotypic (e.g., *dangerous*) or stereotype-neutral (e.g., *tiresome*) word) was a word or not. Analysis revealed that participants who contemplated a past success with egalitarianism responded faster to Black-stereotypic words following Black than White faces; this pattern is indirectly indicative of stereotype activation and is not consistent with enacting an egalitarian goal. In contrast, participants who contemplated a past failure with egalitarianism responded equally fast to Black-stereotypic words following Black and White faces; this pattern is indirectly indicative of stereotype control and consistent with enacting an egalitarian goal. In a replication using a similar LDT (where participants were directed to ignore rather than memorise faces), Moskowitz and Li (Experiment 4) demonstrated that participants who contemplated a past failure in relation to egalitarianism responded significantly slower to Black-stereotypic words following Black than White faces; this pattern is indirectly indicative of stereotype inhibition and consistent with enacting an egalitarian goal. In sum, Moskowitz and Li have demonstrated that contemplating success leads to a lack of goal-consistent behaviour (i.e., stereotype activation), while contemplating failure leads to goal-consistent behaviour (i.e., lack of stereotype activation or stereotype inhibition).

These findings (Moskowitz & Li, 2011; see also Moskowitz & Stone, 2012) are consistent with Fishbach and colleagues' model. Specifically, Fishbach and colleagues' model posits that when people with an internal goal perceive a small goal–behaviour discrepancy they exhibit decreased goal adherence. However, this evidence is indirect: Moskowitz and Li, in asking participants to rate the importance of egalitarianism, did not ask about the source of the egalitarian goal. Although it seems reasonable to suggest that participants in Moskowitz and Li's research were reflecting on personally chosen, internal goals, this remains only

speculative. An aim of the present experiment (Experiment 1) was to prime internal reasons for acting egalitarianism to determine the effect of internal motivation on stereotype accessibility following successful egalitarian goal pursuit, providing a more direct test of Fishbach and colleagues' model.

1.3 Not So Ironic Effect of Success for External Goals

Fishbach and colleagues' model (Fishbach & Dhar, 2005, 2007; Fishbach et al., 2009; Koo & Fishbach, 2008; see Figure 1) proposes that because commitment to external goals is uncertain, individuals who are considering external goals question their commitment. Through a process of self-perception, perceiving a small goal–behaviour discrepancy signals that commitment is high and results in increased goal adherence. As a result, perceiving a small goal–behaviour discrepancy in relation to an external egalitarian goal may result in *control* over stereotype activation and prejudice.

A very different prediction, however, can be derived from Moskowitz and Li's (2011; see also Moskowitz, 2002) theoretical reasoning about the impact of success on egalitarian goal pursuit. They argue that while thinking about either a past success or failure in relation to egalitarianism increases the accessibility of the concept "egalitarianism" (i.e., semantic activation); only thinking about a past failure increases the accessibility of the egalitarian *goal*. Specifically, when contemplating a past failure in relation to egalitarianism, people experience psychological tension and pursue the egalitarian goal in order to resolve the tension and meet the desired standard. Conversely, contemplating a past success in relation to egalitarianism causes the psychological tension to cease as the very process of thinking about success affirms oneself as an egalitarian person. As a result, accessibility of the egalitarian goal decreases, leaving only semantic activation of the concept "egalitarianism". Based on this reasoning, it could be concluded that people with internal and external egalitarian goals

should respond identically following success; that is, contemplating successful egalitarian goal pursuit should ironically result in subsequent behaviour that is not consistent with the egalitarian goal regardless of motivation.

Whether success does or does not motivate subsequent pursuit of external goals has not been empirically tested within stereotyping literature, but suggestive evidence comes from Monin and colleagues (Monin & Miller, 2001; Effron et al., 2012). They demonstrated that the source of participants' motivation does not moderate the effect of successfully establishing oneself as egalitarian (earning moral credentials) on prejudice expression. For example, Monin and Miller (2001) found that after selecting a Black man (vs. a White man) for a position in a consultancy firm, participants were more likely to favour choosing a White man (vs. Black man) for an officer's position, regardless of internal or external motivation to respond without prejudice. This would suggest that for both internally and externally motivated people, successfully establishing themselves as egalitarian licenses the expression of a pro-White (anti-Black) attitude. This is inconsistent with Fishbach and colleagues' model, which predicts that small goal–behaviour discrepancies will influence people with an internal versus external goal differently, but consistent with Moskowitz and Li's (2011; see also Moskowitz, 2002) reasoning that success ironically satisfies the egalitarian goal and results in a lack of goal-consistent behaviour (i.e., stereotype activation).

The evidence seems less than compelling for Fishbach and colleagues' prediction that a small goal–behaviour discrepancy reduces stereotype activation for individuals with an external egalitarian goal. However, the idea that success can motivate rather than undermine goal pursuit is supported by the goal-gradient hypothesis (see Hull, 1932) and the goal-looms-larger effect (see Förster, Grant, Idson, & Higgins, 2001; Förster, Higgins, & Idson, 1998). These indicate that motivation to achieve a goal increases as the discrepancy between one's

current position and goal attainment decreases (see Hull, 1932; Förster et al., 2001; Förster et al., 1998). For example, Kivetz, Urminsky, and Zheng (2006) conducted a series of field studies and demonstrated that as participants neared a goal end-state their effort increased. For example, participants rated songs more frequently to earn a gift card as they neared goal attainment. The fact that successful goal pursuit can motivate further goal pursuit when goal attainment is near supports Fishbach and colleagues' model for external goals, whereby perceiving a small goal–behaviour discrepancy should motivate goal pursuit.

In conclusion, empirical research is required to determine whether Fishbach and colleagues' model is valid in relation to the impact of external egalitarian goals on stereotype activation. The crucial question relates to whether perceiving a small goal–behaviour discrepancy motivates subsequent goal-consistent behaviour (i.e., stereotype control) for people with external egalitarian goals, as Fishbach and colleagues' model suggests. A second aim of the present experiment (Experiment 1), therefore, was to prime external reasons for acting egalitarianism to determine the effect of an external egalitarian goal on stereotype accessibility following successful egalitarian goal pursuit.

2.0 Experiment 1

Experiment 1 aimed to investigate how the source of an egalitarian goal (i.e., internal vs. external) influences stereotype accessibility after reflecting on a successful egalitarian act. In this experiment and in all following experiments, I focused on accessibility of Black stereotypes. It could be argued that the nature of anti-Black stereotyping and prejudice within the UK is not the same as in the USA. However, there is nonetheless evidence that Black people are still a target of prejudice and stereotyping within the UK (e.g., Equality & Human Rights Commission, 2009; Ministry of Justice, 2010), providing one reason for studying the regulation of Black stereotypes. There is also a small body of research showing that prejudice

toward Black people does exist within the UK (e.g., Lepore & Brown, 1997), a second reason for studying Black people as the target group for stereotyping. Finally, focusing on the accessibility of Black stereotypes also allows for the direct comparison of my experiments with prior research, which has also focused on the accessibility of Black stereotypes when assessing how an internal egalitarian goal influences stereotype accessibility after reflecting on a past successful or failed egalitarian act (e.g., Moskowitz & Li, 2011; Moskowitz & Stone, 2012).

In Experiment 1, participants were primed with either internal or external reasons to pursue an egalitarian goal. Next, participants completed a success framing task that utilised self-reflection as a means of highlighting a sense of progress/success in relation to egalitarianism. Specifically, participants recalled a time where they successfully acted egalitarian towards a Black person because they felt either personally compelled to do so (internal goal condition) or socially/politically influenced to do so (external goal condition) (adapted from Moskowitz and colleagues; Moskowitz, 2002; Moskowitz & Li, 2011; Moskowitz et al., 2011; Moskowitz & Stone, 2012). Afterward, participants completed an LDT, whereby facilitated response times to Black-stereotypic (vs. stereotype-neutral) words were taken as indicative of stereotype accessibility.

Fishbach and colleagues' model (see Figure 1) proposes that because people with internal goals are committed to their goals, they focus on their goal progress. Perceiving a small goal–behaviour discrepancy should signal that sufficient progress has been made towards the goal, licensing the disengagement from goal pursuit. I, therefore, predicted that when successful goal pursuit is salient, participants in the internal goal condition should exhibit decreased goal adherence by demonstrating stereotype accessibility. Fishbach and colleagues' model also proposes that because people with external goals are uncertainly

committed to their goals, they focus on their goal commitment. Perceiving a small goal–behaviour discrepancy should signal that commitment is high, increasing goal adherence. I, therefore, predicted that when successful goal pursuit is salient, participants in the external goal condition should exhibit increased goal adherence by demonstrating little to no stereotype accessibility.

Additionally, I investigated two potential moderators. First, I examined participants' identification with the external reference group imposing the external egalitarian goal. Research by Fishbach, Henderson, and Koo (2011; see also Sassenberg, Matschke, & Scholl, 2011) suggests that individuals who identify strongly with their group "take on" the group's goals, and consequently view the group's goals as internally generated rather than externally imposed. I, therefore, predicted that identification would moderate the effects of goal source on stereotype accessibility, with participants low in identification demonstrating the pattern predicted by Fishbach and colleagues' model for external goals, but participants high in identification treating the goal as internally generated even when framed as externally imposed by the group.

Second, I investigated goal importance as a potential moderator. According to Locke and Latham (2005), when a goal is important to a person, the likelihood of acceptance increases, certain commitment is induced, and persistent goal-congruent behaviour is likely to occur. Therefore, I predicted that participants who consider the egalitarian goal to be important might respond consistently with the goal by controlling stereotypes regardless of goal source.

Additionally, I included two variables to test aspects of Fishbach and colleagues' model. First, Fishbach and colleagues' model indicates that commitment is certain for people with internal goals, but uncertain for people with external goals. As a result, goal commitment

should be significantly higher in the internal versus the external goal condition. Second, Fishbach and colleagues' model suggests that goal pursuit relaxes following the detection of a small goal–behaviour discrepancy for people with internal goals so that these individuals can pursue other, even contradictory, goals. As positive distinctiveness is incompatible with acting egalitarian, this may serve as a competing goal that internally motivated people seek when they perceive a small goal–behaviour discrepancy. I predicted that participants in the internal goal condition who held a competing positive distinctiveness goal would exhibit greater stereotype accessibility than those who did not hold a competing positive distinctiveness goal because the increased stereotype accessibility would aid the competing positive distinctiveness goal. Consequently, two measures of positive distinctiveness were included in Experiment 1: one assessing the perceived overlap between the White ethnic group and the Black ethnic group (Turner & Crisp, 2010), and one assessing how positively the White ethnic group is viewed in comparison to the Black ethnic group.

2.1 Method

2.1.1 Participants and design. One-hundred White students[1] from the University of Birmingham completed the experiment in exchange for course credit or money (£3). One participant withdrew from the experiment, and 33 participants' data (18 in the internal goal condition, 15 in the external goal condition) were removed from data analysis for failing to describe a successful egalitarian act that they had actually committed; in a follow-up question, these participants reported having described an imagined rather than actual act of

[1] Based on past research by Fishbach et al. (2011), I expected to achieve a cohen's *d* somewhere between 0.6 to 0.7 when examining the moderating effect of identification on the influence of goal source on stereotype accessibility. Aprior power analyses (G*Power; Faul et al., 2009) using effect size F^2 (0.09 and 0.13; determined using DeCoster's (2012) effect size conversion calculator), critical alpha (.05), total sample size (119 participants), and number of predictors (2), indicated that 40–60 participants per condition would be required to achieve adequate power at 0.80. In Experiment 1, I collected 50 participants per condition.

egalitarianism. I chose to remove these participants from the analyses on the grounds that if they were unable to recall an actual experience, then they may actually have experienced themselves as failing rather than succeeding in acting egalitarian. This left 66 participants (56 female; M_{age} = 19.4 years, SD = 1.14) who were assigned to one of two conditions of a single-factor (Goal Source: internal vs. external) between-participants design.

2.1.2 Materials and procedure. A female experimenter greeted groups of one to four participants and informed them that the experiment aimed to determine the cognitive processes behind mental and visual imagery. Subsequently, participants sat in individual cubicles, in front of personal computers running MediaLab and DirectRT research software (Empirisoft Corporation, 2008).

2.1.2.1 Instructional manipulation check A. To increase participants' diligence, participants completed an instructional manipulation check at the beginning of the experiment (Oppenheimer, Meyvis, & Davidenko, 2009). As depicted in Figure 2, participants indicated the type of computer they were using to complete the experiment (in reality, always a standard desktop computer). However, the instructions asked participants to answer incorrectly by selecting "laptop" to indicate that they had read the instructions. Participants who responded incorrectly were asked to "read the instructions carefully in order to continue" and answer the question again. Oppenheimer et al. (2009) has demonstrated that forcing non-diligent participants to read the instructions leads them to respond identically to diligent participants on subsequent tasks. In the present experiment, 23% of 66 participants failed the instructional manipulation check A (12% failed once, 5% failed twice, 3% failed three times, and 3% failed four or more times) before correctly responding and continuing the experiment.

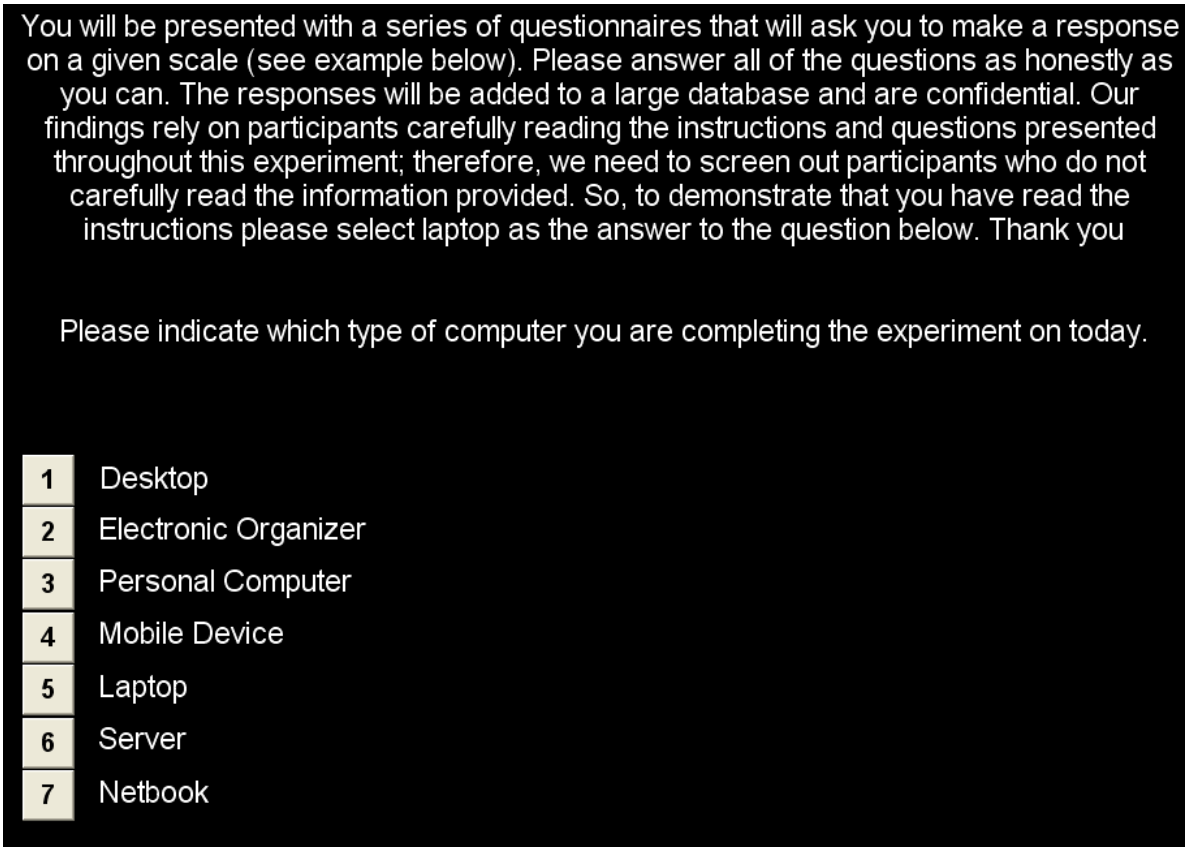

Figure 2. **Screen capture of the instructional manipulation check A.**

2.1.2.2 Goal source manipulation. Next, participants were primed with either an internal or external egalitarian goal. Specifically, half of participants rated their agreement with 10 internal reasons for acting egalitarian towards Black people (e.g., "I personally value being fair towards Black people"; see Appendix B), presented in a random order. The other half of the participants rated their agreement with 10 external reasons to act egalitarian towards Black people (e.g., "People believe that others should treat Black people equally"; see Appendix B), presented in a random order. Participants rated their agreement on a 7-point scale ranging from -3 (*disagree strongly*) to +3 (*agree strongly*). This internal/external goal source manipulation was adapted from Plant and Devine's (1998) internal and external motivation to control prejudice scales.

2.1.2.3 Success framing task. Next, participants learned that their memory-based imagery would be assessed. Participants spent three minutes describing a personal experience (minimum of five lines) in which they successfully acted in an egalitarian manner towards a Black person because they felt personally compelled or socially/politically influenced to do so; participants described a personal experience that matched the goal-source manipulation to which they were assigned for the previous task (i.e., participants who were primed with an internal goal wrote about a personally-driven act, whereas participants who were primed with an external goal wrote about an act that was prompted by others' standards). Participants were instructed to provide as much detail as possible, as the quantity and quality of their imagery would be assessed. Furthermore, participants were also provided with a definition of egalitarianism to assist them with the task: "Egalitarianism means acting fair, just, and tolerant of others as well as treating people equally regardless of whether they differ from you and regardless of their ethnicity, religious background, gender, sexual orientation, physical appearance, etc." (based on Moskowitz, 2002, p. 401). This goal progress framing task was a modified version of the essay-writing task used by Moskowitz and colleagues to manipulate participants' focus on successful and failed goal pursuit (Moskowitz, 2002; Moskowitz & Li, 2011; Moskowitz et al., 2011; Moskowitz & Stone, 2012).

2.1.2.4 Lexical decision task. Next, participants completed an LDT to assess their activation of the Black stereotype. Participants learned that the cognitive processes behind the mental imagery of words would be assessed. Participants determined as quickly and accurately as possible whether a string of letters made up a word in the English language, using keys labelled "yes" and "no". First, participants completed eight practice trials (including four neutral-words and four non-words) to acquaint themselves with the task. Subsequently, participants completed two blocks of 56 trials; the blocks contained the same

stimuli, presented in random order: 28 words equated for valence and word length (see Appendix C), comprising 14 Black-stereotypic words (seven positive, e.g., *athletic*; seven negative, e.g., *loud*; taken from prior research (e.g., Devine, 1989; Devine & Elliot, 1995; Lepore & Brown, 1997; Livingston & Brewer, 2002; Kawakami et al., 2000; Sassenberg & Moskowitz, 2005; Wheeler & Fiske, 2005; Wittenbrink et al., 1997))[2], 14 stereotype-neutral words (seven positive, e.g., *pleasant*; seven negative, e.g., *disgust*; adopted from Livingston & Brewer, 2002), and 28 non-words, created by scrambling the word stimuli, with the constraint that the non-words were pronounceable (e.g., *loud* was scrambled as *olud*). On each trial, a fixation cross was presented for 500 ms, followed by a letter string that remained onscreen until participants made a response. If participants responded faster than 150 ms, a message appeared instructing participants to wait for the letter string to appear; if participants responded slower than 1750 ms, a message appeared instructing participants to respond faster. The intertrial interval was 500 ms.

2.1.2.5 Instructional manipulation check B. A second instructional manipulation check was included to determine whether participants were still diligent when reading instructions. As depicted in Figure 3, participants indicated how many experiments they had participated in that day. This would typically be a low number; however, the instructions asked participants to answer incorrectly by selecting "eleven or more" to indicate that they had read the instructions. In total, 11% of 66 participants failed the instructional manipulation test B once before correctly responding and continuing the experiment.

[2] A subset of the Black-stereotypic words has been used in a UK context (see Lepore & Brown, 1997). The remaining items have been used in both a USA and a European context (see Devine, 1989; Devine & Elliot, 1995; Livingston & Brewer, 2002; Kawakami et al., 2000; Sassenberg & Moskowitz, 2005; Wheeler & Fiske, 2005; Wittenbrink et al., 1997).

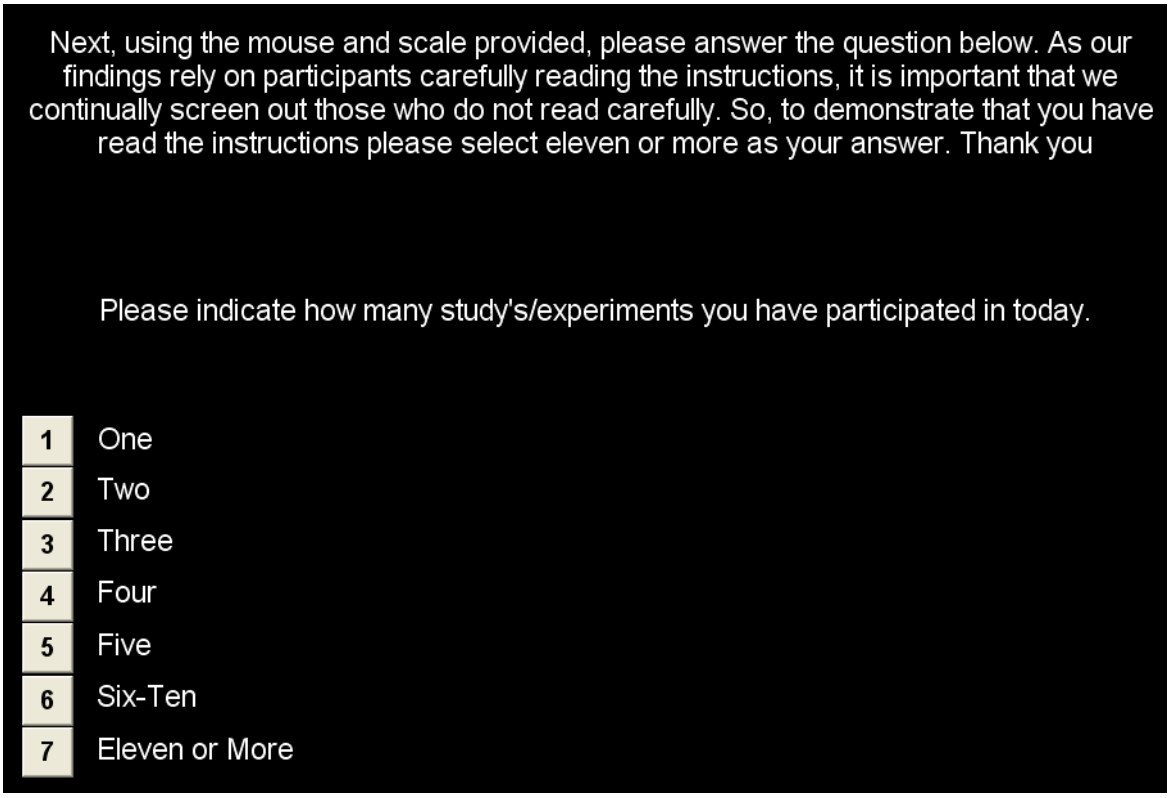

Figure 3. **Screen capture of the instructional manipulation check B**

2.1.2.6 Actual–ideal differentiation. Next, participants completed Turner and Crisp's (2010) actual–ideal differentiation measure. Participants indicated what they thought the actual and the ideal relationship between the White ethnic group and the Black ethnic group was. Participants achieved this by selecting one of seven pictures depicting varying degrees of overlap between the two groups twice, once for the actual relationship and once for the ideal relationship (see Figure 4). The order of presentation of the actual and ideal question versions was randomised.

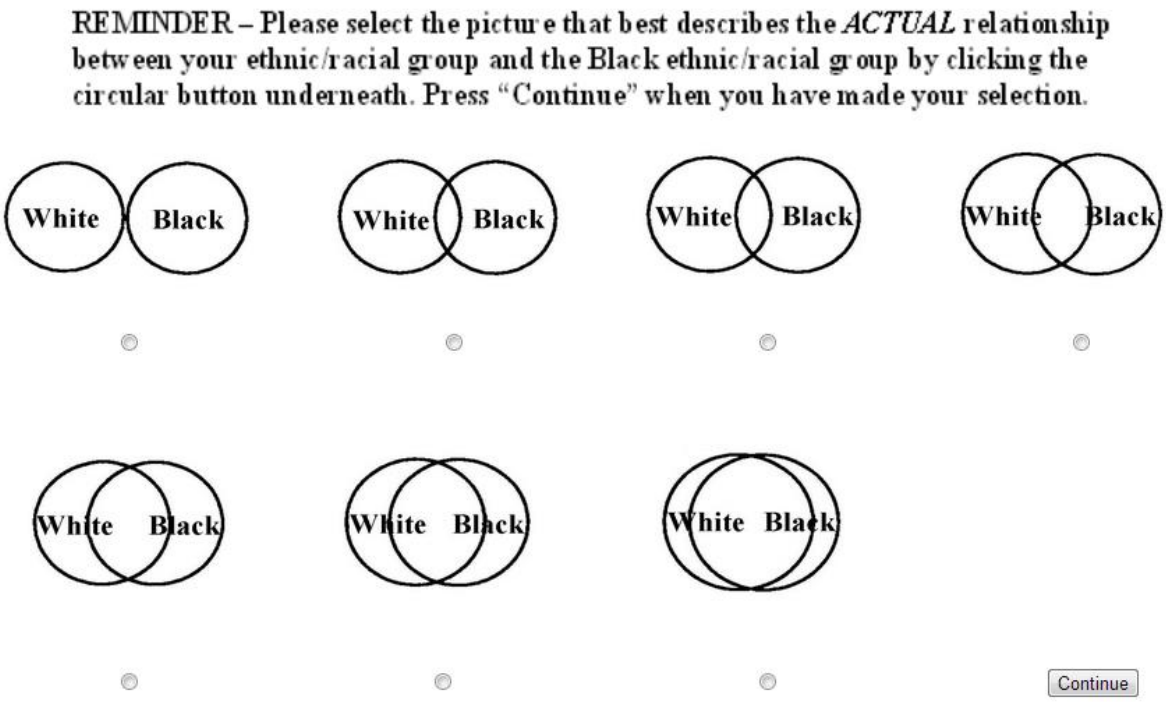

Figure 4. **Screen capture of the actual differentiation question.** The ideal differentiation version of the question was created by replacing "ACTUAL" with "IDEAL".

2.1.2.7 Goal commitment. Next, participants rated how much they agreed or disagreed with a single item assessing their goal commitment ("I am strongly committed to pursuing the goal of being egalitarian (i.e., fair, tolerant, and equal) towards Black people"), on a 7-point scale ranging from -3 (*disagree strongly*) to +3 (*agree strongly*).

2.1.2.8 Identification with White ethnic group. Next, participants rated how much they agreed or disagreed with 5 randomly ordered statements, on a 7-point scale ranging from -3 (*disagree strongly*) to +3 (*agree strongly*): four items from Crisp and Beck (2005; e.g., "I identify strongly with members of my ethnic group") and one item from Badea, Jetten, Czukor, and Askevis-Leherpeux (2010; "I perceive myself as being similar to other members of my ethnic group").

2.1.2.9 Goal importance. Next, participants rated the personal importance of three randomly presented egalitarian behaviours (e.g., "being tolerant of Black people", "acting fairly towards Black people", and "treating Black people equally"), on a 6-point scale ranging from 0 (*not important*) to 5 (*extremely important*).

2.1.2.10 Manipulation check for success framing task. Next, participants indicated whether they had actually experienced the personal experience they described in the success framing task.

2.1.2.11 Positive distinctiveness questions. Finally, participants rated how positively their ethnic group is viewed relative to the Black ethnic group according to other members of their ethnic group ("How positive is your ethnic group relative to the Black ethnic group, as viewed by other members of your ethnic group?"), and how positively their ethnic group is viewed ("How positively is your group viewed?"). Participants rated each question in a 5-point scale ranging from 0 (*not at all positive*) to 4 (*completely positive*).

Upon completing these measures, participants provided demographic information before being provided standardised debriefing information and thanked for their participation.

2.2 Results

2.2.1 Manipulation checks.

2.2.1.1 Goal source primes. A goal prime index was created by averaging responses, with higher scores indicating greater agreement with the goal primes. Participants in the internal goal condition agreed moderately with the goal primes (M = 2.29 on a -3 to +3 scale, SD = 0.78, α = .87); participants in the external goal condition agreed somewhat to moderately with the prime statements (M = 1.50, SD = 0.56, α = .51).

2.2.2 Stereotype accessibility[3]. The main dependent variable was mean reaction times (RTs) to Black-stereotypic and stereotype-neutral words on the LDT. One participant's data were excluded from the analyses as this participant made incorrect responses on more than 15% of trials. For the remaining 65 participants, trials with incorrect responses (4.45% of the data) and reaction times exceeding 2.5 standard deviations away from each participant's individual mean RT (2.90% of the data) were excluded from the analyses.

The data were submitted to a 2 (Goal Source: internal vs. external) × 2 (Word Type: Black-stereotypic vs. stereotype-neutral) × 2 (Word Valence: negative vs. positive) × 2 (Block: 1 vs. 2) mixed-model analysis of variance (ANOVA) with goal source as a between-participants factor. The analysis revealed two significant main effects. First, a significant main effect of block, $F(1, 63) = 13.27$, $p = .001$, $\eta^2_p = .17$, indicated that participants responded faster in Block 2 (M = 488 ms, SE = 7.04) than Block 1(M = 504 ms, SE = 8.40). Second, a significant main effect of word valence, $F(1, 63) = 34.90$, $p < .001$, $\eta^2_p = .36$, indicated that participants responded faster to positive words (M = 486 ms, SE = 7.16) than negative words (M = 507 ms, SE = 8.04).

The analysis also revealed two significant interactions. First, there was a significant Word Type × Word Valence interaction, $F(1, 63) = 9.27$, $p = .003$, $\eta^2_p = .13$. For positive words, participants responded equally fast to Black-stereotypic words (M = 488 ms, SE = 7.53) and stereotype-neutral words (M = 483 ms, SE = 7.26), $t(64) = 1.40$, $p = .17$, $d = 0.09$. For negative words, participants responded faster to Black-stereotypic words (M = 499 ms, SE

[3] When conducting the stereotype accessibility analyses with the entire sample, regardless of whether or not participants described an actual or imagined successful egalitarian act, or with participants who described an imagined rather than actual successful egalitarian act, participants in both the internal and external goal conditions exhibited accessibility of negative Black stereotypes.

= 7.54) than stereotype-neutral words (*M* = 513 ms, *SE* = 8.94), *t*(64) = 3.04, *p* = .003, *d* = 0.21.

Second, and more importantly, the analysis yielded a significant Goal Source × Word Type × Word Valence interaction, *F*(1, 63) = 5.18, *p* = .026, η^2_p = .08. Interaction means are presented in Figure 5. The interaction was decomposed by analysing the internal and external goal conditions separately.

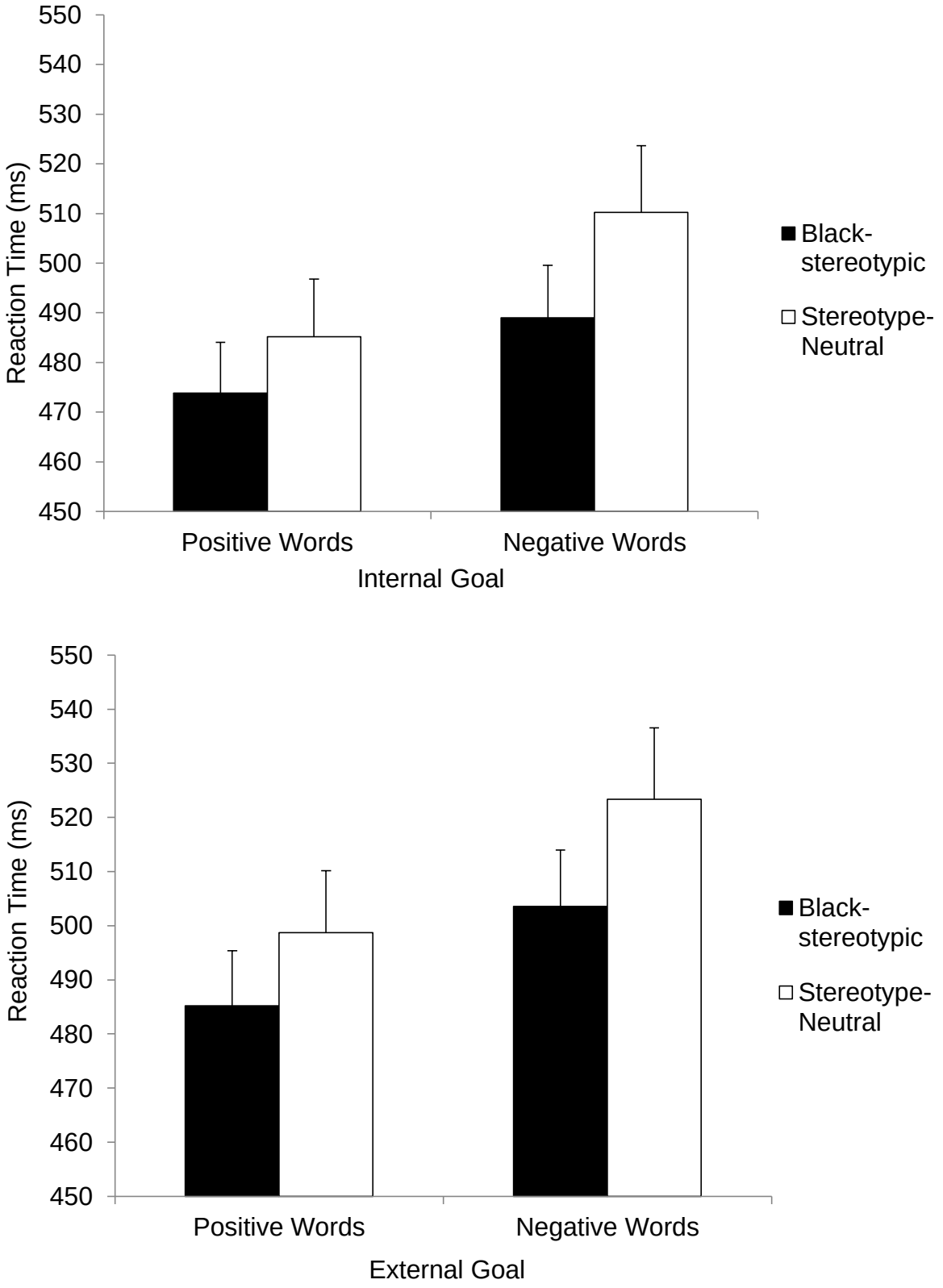

Figure 5. **Mean reaction times (ms) as a function of goal source, word valence, and word type, Experiment 1.** *Note.* Error bars represent standard error.

2.2.2.1 Internal goal condition. A 2 (Word Type: Black-stereotypic vs. stereotype-neutral) × 2 (Word Valence: negative vs. positive) repeated-measures ANOVA revealed two significant main effects. First, a significant main effect of word valence, $F(1, 31) = 12.87$, $p = .001$, $\eta^2_p = .29$, indicated that participants responded faster to positive words (M = 479 ms, SE = 10.39) than negative words (M = 499 ms, SE = 11.86). Second, a significant main effect of word type, $F(1, 31) = 4.13$, $p = .051$, $\eta^2_p = .12$, indicated that participants responded faster to Black-stereotypic words (M = 485 ms, SE = 10.27) than stereotype-neutral words (M = 494 ms, SE = 11.71).

The analysis also revealed a significant Word Type × Word Valence interaction, $F(1, 31) = 9.56$, $p = .004$, $\eta^2_p = .24$. For positive words, participants responded equally fast to Black-stereotypic words (M = 483 ms, SE = 11.08) and stereotype-neutral words (M = 475 ms, SE = 10.59), $t(31) = 1.32$, $p = .20$, $d = 0.13$. For negative words, however, participants responded faster to Black-stereotypic words (M = 486 ms, SE = 10.17) than stereotype-neutral words (M = 512 ms, SE = 14.36), $t(31) = 3.37$, $p = .002$, $d = 0.36$.

2.2.2.2 External goal condition. A 2 (Word Type: Black-stereotypic vs. stereotype-neutral) × 2 (Word Valence: negative vs. positive) repeated-measures ANOVA revealed a significant main effect of word valence, $F(1, 32) = 25.47$, $p < .001$, $\eta^2_p = .44$, such that participants responded faster to positive words (M = 492 ms, SE = 9.89) than negative words (M = 513 ms, SE = 10.63). No additional main effects or interactions were significant, all $p > .40$.

2.2.3 Potential moderators. A negative stereotype accessibility index[4] was created to explore the potential moderating role of identification with the external reference group (i.e.,

[4] I decided on a per-experiment basis whether to include both positive and negative words when creating the stereotype accessibility index. If the main stereotype accessibility analysis

White ethnic group in the present experiment), and of goal importance, on the influence of goal source on stereotype accessibility. Specifically, response times to negative Black-stereotypic words were subtracted from response times to negative stereotype-neutral words, with higher scores indicating stronger negative stereotype accessibility (*Grand M* = 28 ms, *SD* = 76.62).

2.2.3.1 Identification with White ethnic group. An index of identification was created by averaging responses, with higher scores indicating higher identification with the White ethnic group (*M* = 1.16, *SD* = 1.00, α = .86[5]); as the mean score indicates, the majority of participants agreed somewhat with the identification items. The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Identification standardised regression analysis. The analysis revealed a significant main effect of goal source, $\beta = 23.31$, $t(61) = 2.55$, $p = .013$, $R^2 = .12$; accessibility of Black stereotypes was greater in the internal than the external goal condition. No additional main effects or interactions were significant, all $p > .15$; identification had no effect.

2.2.3.2 Goal importance. An index of goal importance was created by averaging across the three items, with higher scores indicating higher goal importance (*M* = 3.63, *SD* = 0.61, α = .90). The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Goal Importance standardised

---

included no word-valence effects, then I included both positive and negative words. In the present experiment, I confined the stereotype accessibility index to negative words because in the main stereotype accessibility analysis, significant effects only emerged for negative words.

[5] One of the items in the identification measure ("I perceive myself as being similar to other members of my ethnic group") could be considered more a measure of prototypicality (i.e., how representative of the group the participant feels they are) than of identification (i.e., the extent to which the participant has adopted the identity of the group and sees that identity as emotionally significant; Tajfel & Turner, 1979). However, removing this item from the identification index actually reduced the reliability of the index (from α = .86 to α = .80). Consequently, this item was retained in the identification index.

regression analysis. The analysis revealed a significant main effect of goal source, $\beta = 22.09$, $t(61) = 2.33$, $p = .023$, $R^2 = .11$; accessibility of Black stereotypes was greater in the internal than the external goal condition. No main effects or interactions were significant, all $p > .31$; goal importance had no effect.

2.2.3.3 Positive distinctiveness. Due to a programming error, data from the actual–ideal differentiation measure was lost. The two positive distinctiveness items were averaged into an index ($M = 2.33$, $SD = 0.71$; $\alpha = .79$). The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Positive Distinctiveness standardised regression analysis. The analysis revealed a significant main effect of goal source, $\beta = 22.85$, $t(61) = 2.46$, $p = .017$, $R^2 = .093$; accessibility of Black stereotypes was greater in the internal than the external goal condition. No main effects or interactions were significant, all $p > .75$; positive distinctiveness had no effect.

2.2.4 Additional analyses.

2.2.4.1 Goal commitment. Overall, goal commitment was high ($M = 2.41$, $SD = 0.88$). A one-way between-participants ANOVA indicated that goal commitment did not differ between the internal and external goal conditions, $F(1, 64) = 2.44$, $p = .12$, $\eta^2_p = .04$.

3.0 Discussion

The findings of Experiment 1 demonstrated that after reflecting on a successful egalitarian act, participants primed with internal reasons to act egalitarian exhibited accessibility of negative Black stereotypes, whereas participants primed with external reasons to act egalitarian exhibited little to no accessibility of Black stereotypes. This pattern suggests that after reflecting on a past success, participants in the external goal condition controlled their stereotypes while participants in the internal goal condition did not. This pattern is consistent with my hypotheses, derived from Fishbach and colleagues' model (see Figure 1).

Fishbach and colleagues' propose that people with internally generated goals focus on goal progress and that evidence of goal achievement leads these individuals to "relax" their focus on the egalitarian goal, thereby decreasing goal adherence. In the current experiment, providing participants in the internal goal condition with evidence for goal achievement led to a decrease in goal adherence (in this case, stereotype accessibility). In contrast, Fishbach and colleagues' propose that people with externally imposed goals focus on goal commitment and that evidence of goal achievement leads these individuals to feel highly committed to the goal, thereby increasing goal adherence. In the current experiment, providing participants in the external goal condition with evidence for goal achievement led to an increase in goal adherence (in this case, little to no stereotype accessibility).

To the best of my knowledge, this is the first experiment to examine the effect of an external goal on stereotype accessibility after reflecting on a past success. Past research by Moskowitz and colleagues (Moskowitz & Li, 2011; Moskowitz & Stone, 2012) examining the effect of an *internal* goal on stereotype accessibility after contemplating success, demonstrated that people with an internal goal to act egalitarian exhibited stereotype activation after contemplating success but stereotype control after contemplating failure. The results of the present experiment, therefore, are also consistent with Moskowitz and Li's (2011; see also Moskowitz, 2002) reasoning for *internal* goals in relation to success. More importantly, however, the present experiment indicated that contemplating success has the opposite effect for people who have an *external* goal to act egalitarian. Rather than exhibiting accessibility of Black stereotypes (suggestive of stereotype activation), as people with an internal egalitarian goal did, people with an external egalitarian goal exhibited little to no accessibility of Black stereotypes (suggestive of stereotype control), consistent with the goal to act egalitarian. This suggests that approaching goal attainment is motivating for externally

motivated participants, which is inconsistent with Moskowitz and Li's theory that contemplating success only results in semantic activation of the concept "egalitarianism" and leads to the ironic activation of stereotypes. If Moskowitz and Li's theory were true, then participants in the internal and external goal condition should have responded identically, with both groups exhibiting stereotype accessibility after reflecting on a past success.

Although past research has provided evidence that identification with the group imposing the external goal is a key moderator of goal adherence (see Fishbach et al., 2011; Sassenberg et al., 2011), the present experiment failed to replicate this effect; that is, identification with the White ethnic group who imposed the external goal did not influence whether stereotype activation occurred. However, the external reference group used in Experiment 1 may not have been appropriate. I chose the White ethnic group as the group advocating for politically correct standards in the priming procedure. According to optimal distinctiveness theory (Brewer, 1991), people are motivated to balance between two competing motivations: the need to belong within a group (assimilation) and the need to be individual (differentiation). Larger groups, like the White ethnic group, allow a large number of people to qualify for inclusion in the group, which may leave group members seeking greater differentiation of the self from the group (Brewer, 1991). In contrast, smaller groups allow people to balance between these competing motivations (Brewer, 1991). Consequently, individuals may identify more strongly with smaller groups. Indeed, Simon and Hamilton (1994) found that participants self-categorised themselves as members of a group (proxy for identification with group) to a greater extent when the group was small compared to large. Consequently, identification with the White ethnic group who imposed the external egalitarian goal may not have influenced whether stereotype activation occurred because the White ethnic group was too broad and inclusive for participants to identify strongly with,

particularly in the context of egalitarianism, as the "White ethnic group" may also be perceived as including nonegalitarian people (e.g., British National Party). Indeed, participants tended to self-report as only somewhat positive with regarding to ethnic identification, which may have obviated any role for identification.

In conclusion, Chapter 2 aimed to investigate the effect of internal and external egalitarian goals on stereotype accessibility following successful goal pursuit. This experiment demonstrated that after contemplating a successful egalitarian act, those who were primed with an external egalitarian goal demonstrated little to no accessibility of Black stereotypes, whereas those primed with an internal egalitarian goal demonstrated accessibility of negative Black stereotypes. To the best of my knowledge, this is the first experiment to examine the effect of an external goal on stereotype accessibility after contemplating success.

CHAPTER 3

CONSEQUENCES OF COMPLIANCE VERSUS NONCOMPLIANCE WITH INTERNAL OR EXTERNAL EGALITARIAN GOALS FOR STEREOTYPE ACCESSIBILITY

In this chapter, I examined the effect of receiving external feedback that highlighted the potential for a goal–behaviour discrepancy on stereotype activation in the context of egalitarianism. Research has examined the effect of goal–behaviour discrepancies on stereotype activation for individuals with internally generated goals to act egalitarian. Yet, to the best of my knowledge, there has been no empirical examination of how goal–behaviour discrepancies influence stereotype accessibility for people with externally imposed egalitarian goals. The work in this chapter, therefore, examined the effect of receiving external feedback that highlighted the potential for a large, small, or no goal–behaviour discrepancy on stereotype activation for individuals with either an internal or external goal to act egalitarian. Specifically, across four experiments, participants were primed with either internal reasons (personal beliefs and values) or external reasons (other people's expectations) to act in an egalitarian manner before receiving false feedback from an "egalitarianism test" that ostensibly showed the potential for a large goal–behaviour discrepancy (Experiments 2 & 5), a small goal–behaviour discrepancy (Experiment 3), or a goal–behaviour match (i.e., no goal–behaviour discrepancy; Experiments 4 & 5). In Experiment 2, highlighting the potential for a large goal–behaviour discrepancy rendered stereotypes accessible for participants in the external, but not the internal, goal condition. In Experiments 3 and 4, highlighting the potential for a small goal–behaviour or a goal–behaviour match rendered negative stereotypes accessible for participants in both the internal and external goal conditions. Finally, in

Experiment 5, all participants exhibited accessibility of negative stereotypes regardless of goal source (i.e., internal vs. external) or goal–behaviour status (i.e., goal–behaviour discrepancy vs. goal–behaviour match). These findings have important implications for our understanding of the impact of goal–behaviour discrepancies on self-regulation for internal and external egalitarian goals.

1.0 General Introduction

When people fail to act in accordance with their goals, a goal–behaviour discrepancy may be perceived. Research has examined the prevalence of goal–behaviour discrepancies in relation to prejudice level (high- vs. low-prejudiced attitude; Devine et al., 1991; Monteith, 1996; Monteith et al., 1993; Zuwerink, Devine, Monteith, & Cook, 1996) and motivation source (internal vs. external motivation; Plant & Devine, 1998). In this research, participants completed a discrepancy measure whereby they indicated separately how they *should* and *would* act in a series of hypothetical scenarios. Example scenarios included feeling uncomfortable if a Black man sits next to you on a bus (Devine et al., 1991) or leaving a restaurant because the waiter is gay (Monteith et al., 1993). This highlights any contradictions between participants' actual behaviour (*would*) and their prejudiced standards (*should*), leading to the perception of a discrepancy. Based on participants *would* and *should* scores, a discrepancy score was calculated (subtracting participants *should* ratings from their *would* ratings for each scenario and then summing the discrepancy scores across scenarios). Collectively, this research has demonstrated that people are more likely to report a discrepancy (i.e., where actual responses (*would*) are more prejudiced than their standards (*should*) allow) than no discrepancy (i.e., where actual responses (*would*) matched what their

standards (*should*) allow[6]). Due to the clear prevalence of goal–behaviour discrepancies, research has investigated the impact of goal–behaviour discrepancies on stereotype activation and prejudice.

1.1. Prejudice Regulation

1.1.1 Internal motivation. Research examining the effect of internal motivation on prejudice following self-regulatory failure and success indicates that Fishbach and colleagues' model may be valid in relation to prejudice. Fehr and Sassenberg (2010) examined the effect of internal motivation on prejudice following a self-regulatory *failure*. In their second experiment, German participants first completed the IMS and EMS scales (Plant & Devine, 1998) and then completed a modified IAT task. In the prejudice-failure condition, participants completed a traditional IAT task comprised of one prejudice-congruent block (German faces paired with positive attributes (e.g., *happy*) and Arab faces paired with negative attributes (e.g., *murder*)) and one prejudice-incongruent block (Arab faces paired with positive attributes and German faces paired with negative attributes); post-task feedback indicated that participants' attitude towards Arabs was more negative than their attitude towards Germans. In the no-failure condition, participants completed a modified IAT task comprised of two prejudice-congruent blocks that made the task subjectively easier to participants; moreover, participants did not receive any feedback to suggest failure. Participants then completed a modified Donald paradigm to assess prejudice; they were asked to read six stories, each one

[6] In Plant and Devine's (1998) research, "no discrepancy" for externally motivated participants only included participants who indicated that their actual responses (*would)* were less prejudiced than their standards (*should*) allow too. For internally motivated participants in Plant and Devine's (1998) research, and for high- and low-prejudiced participants in Monteith and colleagues' research (Devine et al., 1991; Monteith, 1996; Monteith et al., 1993; Zuwerink et al., 1996), those indicating that their actual responses (*would*) were less prejudiced than their standards (*should*) allow were excluded from the analyses or analysed separately.

about a different man completing a series of ambiguous actions, and to form an impression about the man. In half of the stories the man had a German name and in the other half of the stories the man had an Arabic name. Afterward, participants indicated how applicable 10 adjectives were (five positive (e.g., *amiable*) and five negative adjectives (e.g., *aggressive*)). The results demonstrated that internally motivated participants expressed less prejudice (i.e., endorsed fewer negative traits and more positive traits) following a self-regulatory failure compared to no self-regulatory failure.

A recent investigation by Mann and Kawakami (2011) has examined the effect of *successful* goal progress on prejudice for internally motivated individuals. In Experiment 1, while wearing a LifeShirt System T-shirt that monitored their physiological responses, participants viewed images of Black and White faces. Participants were given the goal to evaluate the Black faces positively and received false feedback ostensibly based on their physiological responses about their progress towards the goal. The false feedback was presented in the form of a horizontal bar shaded up to the halfway point, with "Positive Black Evaluations" as an anchor at the one end. Additionally, participants received new feedback after each block (7 in total), presented alongside their previous feedback so they could monitor their progress. Half of the participants received positive feedback after each block (i.e., the shaded area of the bar increased after each block, indicating that they were progressing towards the goal of "Positive Black Evaluations"). The other half of the participants received negative feedback after each block (i.e., the shaded area of the bar decreased after each block, indicating that they were not progressing towards the goal). Afterward, participants took a chair into a cubicle for an interaction with a Black confederate; the confederate estimated the distance between the front of their chair and the participant's on a scale from 1 (*closest possible distance*) to 9 (*furthest possible distance*). Finally, participants

completed an IAT to assess implicit prejudice. Analysis revealed that internally motivated participants who successfully progressed towards the goal sat further away from a Black confederate and exhibited significantly higher IAT scores, indicative of greater implicit prejudice, than participants who failed to progress towards the goal.

1.1.2 High- versus low-prejudiced attitude. Further indirect support for Fishbach and colleagues' model in relation to prejudice is provided by Monteith and colleagues (Monteith 1993; Monteith et al., 2002) who examined the impact of goal–behaviour discrepancies and prejudice level (high- vs. low-prejudiced attitude) on subsequent prejudice regulation. This research has consistently demonstrated that a goal–behaviour discrepancy motivates prejudice regulation in low-prejudiced people only. For example, in Monteith's (1993) first experiment, participants decided whether to accept or reject a law applicant. In the discrepancy-activated condition, the male applicant was gay and the application was designed to be weak so that participants would reject the applicant. In the discrepancy-not-activated condition, the male applicant was heterosexual and the application was designed to be strong so that participants would accept the applicant. After making a decision, participants were informed about the two applicants and told that only sexual orientation differed between the two applications. Participants also learned that people tend to accept the heterosexual applicant but reject the gay applicant. In the discrepancy-activated condition, this information should result in both high- and low-prejudiced participants perceiving a discrepancy between their personal standards and their seemingly prejudiced decision to reject the law applicant[7]. Participants were then asked to provide feedback on an essay summarising the issues that would be covered in a future stereotyping and prejudice workshop. Analysis revealed that

[7] In a pilot test, Monteith (1993) demonstrated that rejecting the law applicant based on sexual orientation was more prejudiced than both high- and low-prejudiced participants' personal standards allowed.

low-prejudiced participants spent longer reading the essay in the discrepancy-activated condition than the discrepancy-not-activated condition, suggesting that low-prejudiced participants were motivated by the discrepancy to regulate prejudice. In contrast, high-prejudiced participants spent similar amounts of time reading the essay in the discrepancy-activated and discrepancy-not-activated conditions, suggesting that high-prejudiced participants were not motivated by the goal–behaviour discrepancy to regulate prejudice.

In a subsequent investigation, Monteith et al. (2002, Experiment 4) examined the impact of goal–behaviour discrepancies arising from low-prejudiced participants' actual behaviour rather than experimental manipulations. Low-prejudiced participants completed an IAT and, afterward, the experimenter pointed out their faster response times to prejudice-congruent than -incongruent trials in order to highlight a goal–behaviour discrepancy. Next, participants indicated whether they liked or disliked names (Black names and White names) and filler items (e.g., animals or objects). The results demonstrated that as the discrepancy between response times to prejudice-congruent and -incongruent trials increased, so too did negative self-directed affect. Furthermore, the guiltier low-prejudiced participants felt about their performance on the IAT, the longer they paused on Black names (i.e., behavioural inhibition) and the more they liked Black names (i.e., nonprejudiced behaviour). Therefore, goal–behaviour discrepancies motivate low-prejudiced participants to regulate prejudice, whether experimentally induced (Monteith, 1993, Experiment 3) or based on actual behaviour (Monteith et al., 2002, Experiment 4). As high-prejudiced people were not examined by Monteith et al. (2002), it remains unclear whether goal–behaviour discrepancies based on actual performance make a difference to high-prejudiced individuals' prejudice regulation.

The research by Monteith and colleagues' (Monteith 1993; Monteith et al., 2002), however, provides only indirect support for Fishbach and colleagues' model. Although it may

be the case that high-prejudiced individuals are more likely to have an external than internal goal with regard to egalitarianism, it is also possible that they could have both internal and external motivations, or that they may react differently to external motivation than moderate- and low-prejudiced people. Indeed, Monteith and colleagues (Devine et al., 1991; Monteith et al., 1993; Zuwerink et al., 1996) have demonstrated that low-prejudiced people's personal standards are more internalised (i.e., highly committed, highly important, and self-defining) than society's standards. However, high-prejudiced people's standards appear to be based on both their personal and society's standards; while high-prejudiced people's personal standards are more internalised than society's standards, this difference is much smaller than for low-prejudiced people, and society's standards are equally or more internalised for high- than low-prejudiced people (Devine et al., 1991; Monteith et al., 1993; Zuwerink et al., 1996). Therefore low-prejudiced people's standards are predominantly based on personal standards, whereas high-prejudiced people's standards are based on a mixture of personal and societal standards.

1.2 Stereotype Regulation

As mentioned in the previous chapter, only research by Moskowitz and colleagues (Moskowitz & Li, 2011; Moskowitz & Stone, 2012) has examined the impact of contemplating an instance of successful or failed egalitarian goal pursuit on stereotype activation for people with an internal egalitarian goal. Moskowitz and colleagues have consistently demonstrated that contemplating a failure to act egalitarian results in stereotype control (indicated by a lack of stereotype activation or by stereotype inhibition) whereas contemplating a successful egalitarian act results in stereotype activation.

1.3 Validity of Fishbach and Colleagues' Model

The evidence discussed in this chapter suggests that Fishbach and colleagues' model may be valid in relation to internal egalitarian goals. Collectively, research has reliably found that individuals with internal egalitarian goals respond consistently with their egalitarian goal by controlling stereotype activation or prejudice following a self-regulatory failure (e.g., Fehr & Sassenberg, 2010; Mann & Kawakami, 2011; Moskowitz & Li, 2011; Moskowitz & Stone, 2012). This is consistent with Fishbach and colleagues' model, which indicates that for internal goals, a large goal–behaviour discrepancy signals that insufficient progress has been made towards the desired end-state, resulting in goal-consistent action in an effort to make progress. Additionally, research has reliably found that individuals with internal egalitarian goals fail to respond consistently with their egalitarian goal by exhibiting stereotype activation or prejudice following self-regulatory success (Mann & Kawakami, 2011; Moskowitz & Li, 2011; Moskowitz & Stone, 2012). This is also consistent with Fishbach and colleagues' model, which indicates that for internal goals, a small goal–behaviour discrepancy signals that sufficient progress has been made, reducing goal adherence as people "relax" their goal pursuit in favour of other, even contradictory, goals that have been neglected.

However, two unanswered questions in relation to external egalitarian goals remain. First, due to the lack of empirical research investigating external egalitarian goals and goal–behaviour discrepancies in relation to stereotype activation and prejudice, it is unclear whether Fishbach and colleagues' model is valid in relation to external egalitarian goals. In relation to prejudice, Fehr and Sassenberg (2010; Experiment 2) conducted the only investigation of how self-regulatory failure impacts prejudice for people who are externally motivated to act nonprejudiced. Specifically, Fehr and Sassenberg found that prejudice did not differ depending on whether externally motivated participants experienced a self-

regulatory failure or no-failure. Fehr and Sassenberg used a *subtle* explicit measure of prejudice, so the lack of significant finding may not be surprising considering that Plant and Devine (1998) demonstrated a weak relationship between external motivation and explicit prejudice measures. Furthermore, no empirical research has examined how small goal–behaviour discrepancies affect prejudice for people who are externally motivated to act egalitarian.

In relation to stereotyping, there has been no investigation of the impact of external egalitarian goals following either self-regulatory success or failure on stereotype activation. Research has shown that stereotyping and prejudice are only moderately correlated (Amodio & Devine, 2006; Talaska, Fiske, & Chaiken, 2008; but see Gawronski, Peters, Brochu, & Strack, 2008), making it important to determine whether goal source and goal–behaviour discrepancies influence stereotype activation differently. As the empirical findings for internal egalitarian goals following self-regulatory success and failure in relation to stereotype activation match Fishbach and colleagues' model (see Moskowitz & Li, 2011; Moskowitz & Stone, 2012), the empirical findings for external egalitarian goals following self-regulatory success and failure in relation to stereotype activation may also match Fishbach and colleagues' model. Additionally, in Experiment 1 of the present thesis, the effect of contemplating a self-regulatory success in relation to egalitarianism resulted in little to no stereotype accessibility only for participants with an external egalitarian goal. This finding provides preliminary evidence that Fishbach and colleagues' model for external goals and small goal–behaviour discrepancies (large goal–behaviour discrepancies remain untested) may be valid in relation to stereotype activation. Therefore, the research in the present chapter aimed to examine how goal–behaviour discrepancies influence stereotype accessibility for internal and, more importantly, external egalitarian goals.

Second, the vast majority of previous research reviewed in this chapter has *not* investigated whether the *magnitude* of the goal–behaviour discrepancy matters for stereotype activation and prejudice. Instead, this research has examined how perceiving a goal–behaviour discrepancy or no goal–behaviour discrepancy influences prejudice (Monteith 1993; Monteith et al., 2002), or how failed or successful egalitarian goal pursuit influences prejudice (Fehr & Sassenberg, 2010; Mann & Kawakami, 2011) and stereotype activation (Moskowitz & Li, 2011; Moskowitz & Stone, 2012). Only one experiment by Monteith et al. (2002, Experiment 4) has examined how the magnitude of low-prejudiced participants' discrepancies (i.e., faster response times to prejudice-congruent than –incongruent trials) on the IAT influenced prejudice. This research indicated that as the size of the discrepancy increased, so too did negative self-directed affect. Furthermore, the guiltier low-prejudiced participants felt, the longer participants attended to Black names and the more participants liked Black names. However, discrepancies were analysed as a continuous variable. Plant and Devine's (1998) and Monteith and colleagues' (Devine et al., 1991; Monteith, 1993; Monteith et al., 1993; Monteith & Voils, 1998; Zuwerink et al., 1996) empirical research demonstrates that small and large goal–behaviour discrepancies have different consequences for *affect* depending on goal source. Plant and Devine (1998) only speculate about how large goal–behaviour discrepancies influence stereotyping and prejudice. It is unclear, therefore, whether small goal–behaviour discrepancies will affect stereotype activation and prejudice differently depending on goal source. Fishbach and colleagues' model predicts that both the source of the goal (internal vs. external) and the magnitude of the goal–behaviour discrepancy have different consequences for motivation and self-regulation. Consequently, the research in the present chapter aimed to manipulate the size of the goal–behaviour discrepancy to more

directly test Fishbach and colleagues' model and determine whether the magnitude of the goal–behaviour discrepancy is important for internal and external egalitarian goals.

1.4 Goal–Behaviour Discrepancies and Affect

Research has examined the influence of prejudice level and goal–behaviour discrepancies on affect. According to Monteith's (Monteith, 1993; Monteith et al., 2002) self-regulation of prejudice (SRP) model, negative self-directed affect is a crucial determinant of whether prejudice regulation occurs. The SRP model proposes that the perception of a goal–behaviour discrepancy threatens low-prejudiced people's nonprejudiced identity and results in negative self-directed affect (e.g., guilt). Negative self-directed affect serves as a punishment and motivates goal-consistent behaviour as a means of reducing the goal–behaviour discrepancy. Only negative self-directed affect motivates goal pursuit; hence, high-prejudiced individuals, who experience negative other-directed affect (e.g., anger) rather than negative self-directed affect, are not motivated to reduce a goal–behaviour discrepancy.

However, empirical evidence for the nature of the influence of prejudiced beliefs and goal–behaviour discrepancies on affect is mixed. Monteith and colleagues' research has demonstrated that low-prejudiced participants report greater negative self-directed affect in response to larger goal–behaviour discrepancies versus smaller goal–behaviour discrepancies (e.g., Devine et al., 1991; Monteith, 1996; Monteith et al., 1993; Monteith & Voils, 1998). In addition, for larger goal–behaviour discrepancies, low-prejudiced participants report greater negative self-directed affect than high-prejudiced participants (e.g., Devine et al., 1991; Monteith, 1993; Monteith et al., 1993; Zuwerink et al., 1996). In contrast, for smaller goal–behaviour discrepancies, negative self-directed affect is low for both low- and high-prejudiced participants (e.g., Devine et al., 1991; Monteith et al., 1993; Zuwerink et al., 1993). This

research appears to indicate that only low-prejudiced people experience compunction, and only when large goal–behaviour discrepancies are perceived.

However, Monteith (1996) questioned whether high-prejudiced people experience no compunction following large goal–behaviour discrepancies or whether the negative self-directed affect that high-prejudiced people experience is merely attenuated in comparison to that reported by low-prejudiced individuals. Indeed, Monteith found a significant positive relationship between goal–behaviour discrepancies and negative self-directed affect for both high- and low-prejudiced people, such that as discrepancies increased, so too did negative self-directed affect. However, the positive relationship was strong for low-prejudiced participants (r = .55), but modest for high-prejudiced participants (r = .35). Therefore, high-prejudiced individuals do experience negative self-directed affect when they perceive a goal–behaviour discrepancy, but to a lesser extent than low-prejudiced individuals.

Aside from negative self-directed affect, Monteith and colleagues have also investigated how goal–behaviour discrepancies influence negative other-directed affect and positive affect for high- and low-prejudiced participants. The findings for both negative other-directed affect and positive affect are inconsistent. For other-directed affect, Monteith et al. (1993) demonstrated that prejudice-level and goal–behaviour discrepancies may interact, with high-prejudiced (but not low-prejudiced) participants exhibiting greater negative other-directed affect in response to larger goal–behaviour discrepancies compared to smaller goal–behaviour discrepancies. Yet, other research by Monteith and colleagues (Devine et al., 1991; Monteith, 1996) has demonstrated only a main effect of prejudice-level; that is, high-prejudiced participants reported greater negative other-directed affect than low-prejudiced participants, irrespective of goal–behaviour discrepancy size. Furthermore, additional research

by Monteith and colleagues (Monteith, 1993; Monteith & Voils, 1998; Zuwerink et al., 1996) has failed to find an effect for negative other-directed affect.

For positive affect Monteith and colleagues (Devine et al., 1991; Monteith, 1996; Monteith et al., 1993) demonstrated that goal–behaviour discrepancies and prejudice level interact, with low-prejudiced (but not high-prejudiced) participants exhibiting lower positive affect in response to larger goal–behaviour discrepancies compared to smaller goal–behaviour discrepancies. Yet, other research by Monteith and colleagues (Monteith, 1993; Zuwerink et al., 1996) has demonstrated only a main effect of goal–behaviour discrepancy; that is, all participants reported lower positive affect in response to larger goal–behaviour discrepancies compared to smaller goal–behaviour discrepancies, irrespective of prejudice-level. Furthermore, additional research by Monteith and Voils (1998) has failed to find an effect for positive affect.

Research that has examined the influence of goal source (i.e., internal and external) and goal–behaviour discrepancies on affect is limited. According to Higgins' (1987) self-discrepancy theory, prejudice-related discrepancies that are based on an individual's personally-generated standards and mandate what a person should do (i.e., internal goal source) result in feelings associated with self-punishment (e.g., guilt and self-contempt). In contrast, prejudice-related discrepancies that are based on other-imposed standards and mandate what a person ought to do (i.e., external goal source) result in feelings associated with punishment from others (e.g., fearful, threatened). Empirical research by Plant and Devine (1998, Experiment 3) supported the predictions made in Higgins' (1987) self-discrepancy theory. Plant and Devine found that internally motivated participants reported greater negative self-directed affect (e.g., guilt) when larger goal–behaviour discrepancies were detected compared to when smaller goal–behaviour discrepancies were detected,

whereas externally motivated participants exhibited low levels of negative self-directed affect regardless of goal–behaviour discrepancies.

However, subsequent research has found mixed support for the predictions made in Higgins' (1987) self-discrepancy theory. For prejudice-related discrepancies based on *personally-generated* standards, Fehr and Sassenberg (2010, Experiment 1) found congruent evidence that internally (versus externally motivated individuals reported higher negative self-directed affect after a self-regulatory failure compared to no-failure or a prejudice-irrelevant failure. However, Monteith, Mark, and Ashburn-Nardo (2010) found incongruent evidence when they conducted an interview study. They found that *both* internally and externally motivated people reported feeling negative self-directed affect after a failure to act egalitarian. For prejudice-related discrepancies based on *other-imposed* standards, Plant and Devine (2001) found congruent evidence that externally motivated participants exhibit greater anger and threatened affect in response to pressure to act nonprejudiced than internally motivated participants. However, research by Fehr and Sassenberg (2010, Experiment 1) found incongruent evidence when they failed to find any difference in other-directed affect for internally or externally motivated individuals after a self-regulatory failure compared to no-failure or a prejudice-irrelevant failure.

Altogether, this research indicates that the source of a nonprejudiced goal and goal–behaviour discrepancies may have important consequences for affect, but that the precise nature of this influence is unclear. The work in the present chapter aimed to examine the influence of an *egalitarian* goal on affect following a small or large goal–behaviour discrepancy. An egalitarian goal is a superordinate goal in that people can choose to pursue multiple sub-goals (e.g., the goal to act nonprejudiced or the goal to avoid stereotyping) in order to achieve the overarching egalitarian goal (see Fishbach, Zhang, & Dhar, 2006).

Therefore, while the work in the present thesis examined the effect of both internal and external goals, and small and large goal–behaviour discrepancies, on automatic stereotype activation, the work in the present thesis could also have examined the effect of both internal and external goals, and small and large goal–behaviour discrepancies on prejudice had a measure of automatic prejudice been used in place of the measure of automatic stereotype activation. The idea that an egalitarian goal could influence both prejudice and stereotyping is supported by Gawronski and colleagues' (Gawronski & Bodenhausen, 2006; Gawronski, Deutsch, Mbirkou, Seibt, & Strack, 2008; but see Amodio & Devine, 2006, for an alternative argument); the researchers theorised that changes in stereotype activation can result in changes in automatic prejudice. For example, Gawronski, Deutsch, Mbirkou, Seibt, & Strack (2008) demonstrated that affirmation training (responding "YES" to stereotype-incongruent name/face–trait pairings and "NO" to stereotype-congruent name/face–trait pairings) reduced stereotype activation *and* negative evaluations.

Furthermore, examining the effect of an *egalitarian* goal on affect following a small or large goal–behaviour discrepancy is important in relation to Fishbach and colleagues' model. Although negative self-directed affect plays a pivotal role in determining subsequent goal pursuit for low-prejudiced individuals in the SRP model (Monteith, 1993; Monteith et al., 2002), Fishbach and colleagues make no reference to affect in their self-regulation model. Yet, affect may be an important part of the mechanism through which goals influence self-regulation. For example, people with an internal goal may assess whether they are satisfied with the size of the goal–behaviour discrepancy by examining how they feel (i.e., whether they are experiencing negative self-directed affect). If they are experiencing negative self-directed affect, they may interpret the goal–behaviour discrepancy as a large discrepancy, whereas if they are not experiencing negative self-directed affect they may interpret the goal–

behaviour discrepancy as a small discrepancy. Examining affect may lead to the extension of Fishbach and colleagues' model.

2.0 The Present Research

The work in the present chapter aimed to examine how the source of an egalitarian goal (internal vs. external) influences stereotype accessibility after perceiving the potential for a goal–behaviour discrepancy in the context of egalitarianism. In Experiment 1 of this thesis (Chapter 2), self-reflection was used a means of making participants feel like they had successfully progressed towards the egalitarian goal. In the present research, however, self-reflection was not employed for two reasons. First, the work in the present chapter aimed to manipulate the size of the goal–behaviour discrepancy in order to test Fishbach and colleagues' model more directly, requiring greater control than could be achieved by allowing participants to generate their own examples. Second, a third of participants were removed from the analyses in Experiment 1 because they were unable to describe a time where they successfully acted egalitarian towards Black people, resulting in undesirable levels of data loss.

In the research reported in this chapter, goal–behaviour discrepancies were induced using external false feedback ostensibly based on participants' performance on a test of egalitarianism. Past research has successfully manipulated goal–behaviour discrepancies using an ostensible assessment of participants' prejudiced behaviour. For example, in Mann and Kawakami's (2011; see also Monteith et al., 2002) first experiment, participants' viewed images of Black and White faces while their physiological responses were measured, and received false feedback (ostensibly based on their physiological responses to these Black faces) via a horizontal bar (the shaded area either increased or decreased after each block to induce a small or large goal–behaviour discrepancy) that induced a sense of progress or

highlighted a lack of progress. Additionally, past research has also used the IAT to induce a goal–behaviour discrepancy (e.g., Fehr & Sassenberg, 2010; Monteith et al., 2002). For example, Fehr and Sassenberg (2010) used the IAT and feedback indicating that participants had exhibited more prejudice towards Arabs than Germans to induce a sense of failure. Therefore, I used modified versions of the IAT within the present experiments to manipulate goal–behaviour discrepancy size.

In Experiment 1, participants reflected on a past successful egalitarian act in order to highlight a sense of progress/success in relation to egalitarianism. In the research reported in this chapter, I gave participants feedback about their performance upon completion of the IAT task. Rather than simply informing participants that their performance on the IAT indicated that they were or were not prejudiced, as in previous research (e.g., Fehr & Sassenberg, 2010), the feedback used in the present experiments visually depicted a discrepancy between participants' performance on the test and their likely future success at attaining the egalitarian goal. This was achieved by presenting participants with a graph that had an arrow extending upwards towards the feedback. The distance between the top of the arrow and maximum score on the graph (100%) was manipulated to induce a small or large goal–behaviour discrepancy. This manipulation of goal–behaviour discrepancy size is similar to Fishbach and colleagues' method of inducing goal–behaviour discrepancies and should, therefore, provide a more direct assessment of whether Fishbach and colleagues' model is valid in the context of stereotype activation.

It is important to note two key points. First, the shift from inducing discrepancies through self-reflection (Experiment 1) to false feedback (Experiments 2–5) should not change the applicability of Fishbach and colleagues' model to stereotype activation. Inducing discrepancies through self-reflection (e.g., contemplating success or failure; e.g., my

Experiment 1; Moskowitz and Li, 2011; Moskowitz & Stone, 2012; Zhang, Fishbach, & Dhar, 2007) or false feedback (e.g., feedback indicating success or failure or feedback highlighting visual discrepancies of different sizes; Fehr & Sassenberg, 2010; Fishbach & Dhar, 2005; Koo & Fishbach, 2008; Mann & Kawakami, 2011) should both result in a pattern of self-regulation that matches Fishbach and colleagues' model. Second, the shift from focusing on past success at being egalitarian to likely future success at attaining the egalitarian goal should not change the applicability of Fishbach and colleagues' model to stereotype activation either. Zhang et al. (2007) demonstrated that discrepancies based on past actions (i.e., the current goal–behaviour discrepancy) and discrepancies based on future actions (e.g., the potential future goal–behaviour discrepancy) do not differentially affect self-regulation. Specifically, Zhang et al. found that focusing on goal progress led to goal disengagement after success, and focusing on goal commitment led to goal congruent action after success, regardless of the source of the discrepancy (i.e., past actions or future actions). However, the effect of discrepancies based on future actions was more exaggerated than the effect of discrepancies based on past actions. Specifically, focusing on goal progress led to more goal disengagement after success, and focusing on goal commitment led to more goal congruent actions after success, when the sense of success arose from future actions compared to past actions.

3.0 Experiment 2

Experiment 2 aimed to investigate how the source of an egalitarian goal influences stereotype accessibility when there is potentially a large goal–behaviour discrepancy in the context of egalitarianism. Participants were primed with either internal reasons or external reasons to pursue an egalitarian goal. Next participants completed an "egalitarianism test" that

ostensibly showed the potential for a large goal–behaviour discrepancy. Afterward, participants completed an LDT to assess stereotype accessibility.

Fishbach and colleagues' model (see Figure 1) proposes that because people with internal goals are committed to their goals, they focus on goal progress. Perceiving a large goal–behaviour discrepancy should signal that insufficient progress has been made towards the goal, resulting in increased goal adherence. I, therefore, predicted that when participants in the internal goal condition become aware of the potential for a large goal–behaviour discrepancy, they should exhibit increased goal adherence by demonstrating little to no stereotype accessibility. Fishbach and colleagues' model also proposes that because people with external goals are uncertainly committed to their goals, they focus on their goal commitment. Perceiving a large goal–behaviour discrepancy should signal that commitment is low, resulting in decreased goal adherence. I, therefore, predicted that when participants in the external goal condition become aware of the potential for a large goal–behaviour discrepancy, they should exhibit decreased goal adherence by demonstrating stereotype accessibility.

The same measures of identification with the external reference group and goal importance that were used in Experiment 1 were also included in Experiment 2. The only change between Experiment 1 and 2 in relation to these measures was that I used egalitarians as the external reference group rather than the White ethnic group. A potential reason for why identification with the group imposing the external goal did not moderate the influence of goal source on stereotype activation in Experiment 1 could be due to the White ethnic group being too broad and inclusive for participants to identify strongly with. In Experiment 2, I chose to use egalitarians as the external reference group as this group should be smaller than the White ethnic group because a significant number of nonegalitarian people, such as the

British National Party, are excluded from the group. The predictions in relation to identification and goal importance remain as outlined in Experiment 1.

The measures of goal commitment and positive distinctiveness that were used in Experiment 1 were also included in Experiment 2. The predictions in relation to goal commitment remain as outlined in Experiment 1. I made a tentative prediction in relation to positive distinctiveness following a large goal–behaviour discrepancy. According to Fishbach and colleagues' model, when internally motivated people perceive a large goal–behaviour discrepancy they should be motivated to pursue the egalitarian goal; therefore, positive distinctiveness should have no moderating role. However, when externally motivated people perceive a large goal–behaviour discrepancy, Fishbach and colleagues' model predicts that the egalitarian goal will be rejected. It seems reasonable to posit that these individuals may subsequently pursue personally important goals that might even contradict the egalitarian goal. As a result, I explored whether participants in the external goal condition who hold a competing positive distinctiveness goal exhibit greater stereotype accessibility than those who do not hold a competing positive distinctiveness goal on the basis that the increased stereotype accessibility might aid a personally held and competing positive distinctiveness goal.

Finally, I also examined the influence of goal source (i.e., internal or external) on affect following a large goal–behaviour discrepancy. Based on past research (Devine et al., 1991; Monteith, 1993; Monteith, 1996; Monteith et al., 1993; Monteith & Voils, 1998; Plant & Devine, 1998; Zuwerink et al., 1993), one might expect to find the following: greater negative self-directed affect reported by participants in the internal goal condition compared to the external goal condition; greater negative other-directed affect reported by participants in the external goal condition compared to the internal goal condition; lower general positive

affect reported by participants in the internal goal condition compared to the external goal condition; and greater general negative affect (e.g., angry and threatened) reported by participants in the external goal condition compared to the internal goal condition. My predictions in relation to affect are tentative for two reasons: First, the findings across past research (Devine et al., 1991; Monteith, 1993; Monteith, 1996; Monteith et al., 1993; Monteith & Voils, 1998; Zuwerink et al., 1993) have been inconsistent. Second, the large majority of research has examined the effect of prejudiced beliefs (i.e., high- vs. low-prejudiced beliefs) rather than goal source (i.e., internal vs. external) on affect. While some researchers argue that any changes in automatic stereotype activation should result in changes in automatic evaluation (Gawronski & Bodenhausen, 2006; Gawronski, Deutsch, Mbirkou, Seibt, & Strack, 2008), other researchers argue that automatic stereotype activation and evaluation are independent constructs (Amodio & Devine, 2006).

3.1 Method

3.1.1 Participants and design. One hundred and nineteen White students[8] (108 female; M_{age} = 18.9 years, SD = 1.28) from the University of Birmingham completed the experiment in exchange for course credit or money (£3). Participants were assigned to one of two conditions of a single-factor (Goal Source: internal vs. external) between-participants design.

3.1.2 Materials and procedure. Participants were seated in individual cubicles, in front of personal computers running MediaLab and DirectRT research software (Empirisoft

[8] Based on past research by Fishbach et al. (2011), I expected to achieve a cohen's *d* somewhere between 0.6 to 0.7 when examining the moderating effect of identification on the influence of goal source on stereotype accessibility. Aprior power analyses (G*Power; Faul et al., 2009) using effect size F^2 (0.09 and 0.13; determined using DeCoster's (2012) effect size conversion calculator), critical alpha (.05), total sample size (119 participants), and number of predictors (2), indicated that 40–60 participants per condition would be required to achieve adequate power at 0.80. In Experiment 2, I collected 60 participants per condition.

Corporation, 2008), and informed that the aim of the experiment was to investigate egalitarianism in university students. Egalitarianism was defined as "acting fair, just and tolerant of others as well as treating people equally regardless of whether they differ from you and regardless of their ethnicity, religious background, gender, sexual orientation, physical appearance, etc."

3.1.2.1 Instructional manipulation check A. The same instructional manipulation check A from Experiment 1 was used in Experiment 2. In total, 38% of 119 participants failed the instructional manipulation check A (18% failed once, 11% failed twice, 3% failed three times, 3% failed four times, and 1% failed five, six, and seven times) before correctly responding and continuing the experiment.

3.1.2.2 Identification with egalitarians. Participants were informed that we wanted to determine how much they identify with egalitarian people before the experiment began. Participants completed a similar measure of identification as in Experiment 1, with two changes. First, the "White ethnic group" was replaced with "egalitarian people" (e.g., "I identify strongly with egalitarian people"). Second, participants indicated how true each statement was using a 9-point scale ranging from 0 (*not at all true*) to 8 (*extremely true*). This measure was completed before the experimental manipulation as a means of highlighting the centrality of egalitarian values to the participants' self-concept.

3.1.2.3 Goal source manipulation. Participants were informed that they would complete several tasks to determine how egalitarian they are, and were ostensibly randomly assigned to focus on Black people to make the tasks less abstract. Participants were then primed with either an internal or external egalitarian goal. Specifically, half of the participants rated their agreement with 10 internal reasons to act egalitarian towards Black people; these statements were identical to those used in Experiment 1 to prime an internal goal. The other

half of the participants rated their agreement with 10 external reasons to act egalitarian towards Black people (e.g., "Other egalitarian people believe that I should treat Black people equally"; see Appendix D), presented in a random order. The external reasons were altered so that the focus was not only on other egalitarian people's beliefs and expectations, but also on egalitarian people's negative reactions to nonegalitarian acts (e.g., "Other egalitarian people would be angry if I treated Black people unequally; see Appendix D). This change was made for two reasons. First, the internal consistency of the external goal primes in Experiment 1 was too low ($\alpha = .51$). Second, participants' agreement with the external goal primes was lower than participants' agreement with the internal goal primes in Experiment 1. Participants rated their agreement on a 7-point scale ranging from -3 (*disagree strongly*) to +3 (*agree strongly*).

3.1.2.4 Large goal–behaviour discrepancy framing task. Next, participants learned that they would be completing a test of how motivated they were to act egalitarian and would receive feedback about their performance, based on how *quickly and accurately* they responded. In reality, participants completed a modified IAT (Greenwald et al., 1998). First, participants completed two practice blocks. On the first practice block, they categorised 10 faces (five Black faces using the "S" key and five White faces using the "K" key) presented in a random order. On the second practice block, participants categorised 10 words (five unpleasant words using the "S" key and five pleasant words using the "K" key; e.g., *disaster* vs. *peace;* see Appendix E), presented in a random order. Then, instead of one block pairing race and valence in a prejudice-congruent manner and one block pairing race and valence in a prejudice-incongruent manner, participants completed two critical blocks that included only prejudice-congruent pairings. On each trial (32 trials in each critical block), participants were asked to categorise Black faces and unpleasant words using the "S" key, and White faces and

pleasant words using the "K" key. The stimuli for the IAT were randomly drawn from a pool of 16 faces (eight Black faces and eight White faces) presented on a black background, and 16 words (eight pleasant and eight unpleasant). Participants also received error feedback in the form of a red "X" each time they made an incorrect response.

The modified IAT in the present experiment was designed to heighten participants' awareness of their own potential for nonegalitarian responding and thus make them more likely to believe failure feedback. This was achieved in three ways. First, participants made only prejudice-congruent categorisations to ensure prejudicial responding was salient. Second, the need for participants to respond quickly and accurately on the test was emphasised; the goal was to inflate error rates. Finally, error feedback was provided to increase participants' awareness of their errors during the test.

3.1.2.5 Experience during the test. After completing the large goal–behaviour discrepancy framing task, participants waited 20 seconds while the computer ostensibly calculated their test score. While waiting to receive their (fictitious) test score, participants answered two randomly ordered questions to assess their experience of the test. Participants indicated how well they thought they had performed on the test, on a 9-point scale ranging from 0 (*not at all well)* to 8 (*extremely well*). Participants also estimated how likely they thought they were to achieve the egalitarian goal based on their performance on the test, on a 10-point scale ranging from *0–10%* to *91–100%.*

3.1.2.6 78% feedback. Given that participants' ratings of their commitment to the egalitarian goal, and the importance of the egalitarian goal were high in Experiment 1, it was likely that participants would tend to view themselves as relatively egalitarian. As a result, 78% was chosen to induce a potentially large goal–behaviour discrepancy to ensure that the feedback was plausible. Using a lower score would have undoubtedly induced a larger goal–

behaviour discrepancy, but receiving feedback that is implausibly low might lead participants to misattribute their performance to the task itself or to discount the test as inaccurate. Indeed, Monteith, Voils, and Ashburn-Nardo (2001) investigated participants' reactions to completing a Black–White IAT, and found that of the participants (64%) who felt the IAT effect (i.e., faster reaction times to prejudice-congruent compared to prejudice-incongruent trials), the majority misattributed their slower responses on the prejudice-incongruent trials as associations with colours (i.e., white and good, black and bad; 26%) or to the task itself (37%). This research indicates that participants tend to misattribute information that contradicts their egalitarian identity to factors that alleviate their accountability.

A pilot test was conducted to determine how participants interpreted 78% (a potentially large goal–behaviour discrepancy) and 98% (a potentially small goal–behaviour discrepancy) feedback. Eight-nine participants were asked to imagine that they had just completed a test that determined how motivated they were to act egalitarian towards Black people. Participants were then told to imagine that the test indicated that they would act egalitarian 78% [98%] of the time, and then asked to indicate how well they felt they performed on the test on a 7-point scale, ranging from -3 (*failed completely*) to +3 (*succeeded completely*). Analysis revealed that participants rated a score of 98% as indicating significantly more successful performance than a score of 78% (1.90 vs. 0.05, respectively), $t(87) = 10.19$, $p < .001$, $d = 1.19$. This demonstrates that participants interpreted 98% as a small goal–behaviour discrepancy, indicating that they had succeeded moderately on the test. However, participants interpreted 78% as a larger discrepancy, indicating neither a success nor a failure.

Although participants responded neutrally to the 78% feedback, it seems reasonable to conclude that this would induce at least a moderate goal–behaviour discrepancy in the present

experiment based upon Sassenberg and colleagues' (2011) reasoning. In their research, participants received false feedback about their studying behaviour (Experiment 2) or their cognitive flexibility (Experiment 3) on a scale from 0% (*inacceptable*) to 50% (*acceptable*) to 100% (*ideal*). Receiving false feedback at 89.2% did not influence positive or negative affect (Experiment 2) or effort to improve the score (Experiment 3). Sassenberg et al. (2011) concluded that receiving 89.2% out of 100% did not surpass participants' high expectations about their performance and consequently was not interpreted as a positive event. Indeed, participants in our pilot study did not interpret 78% favourably (i.e., 78% did not make participants feel even somewhat successful). It is reasonable to conclude, therefore, that receiving false feedback at 78% in the present experiment would have been interpreted as at least a potentially moderate goal–behaviour discrepancy.

Consequently, participants were presented with a graph depicting their performance; an arrow extended upwards to 78% on a scale from 0–100% (see Figure 6). The accompanying wording was consistent with each goal condition.

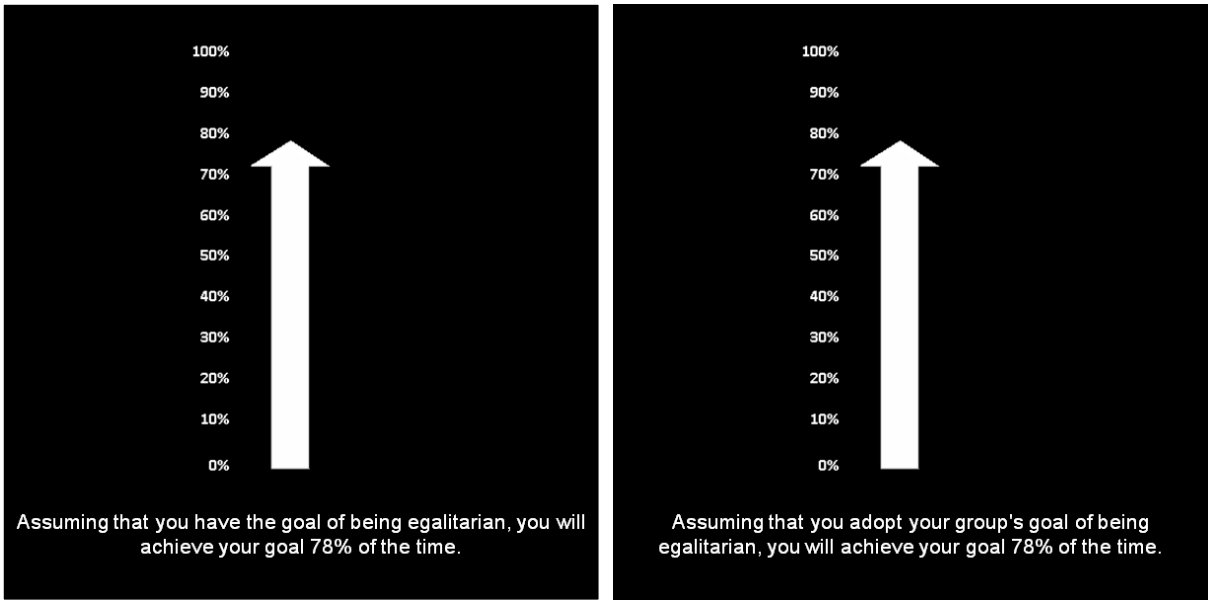

Figure 6. **78% feedback that participants received to induce a potentially large goal–behaviour discrepancy, Experiment 2.** The left panel depicts the feedback in the internal goal condition while the right panel depicts the external goal condition.

3.1.2.7 Affect measures. Immediately after receiving the feedback, participants were presented with the stem sentence, "Please rate the extent to which the following emotions apply to how you are currently feeling". Participants then rated a series of affective items using a 9-point scale, ranging from 0 (*does not apply at all*) to 8 (*applies extremely*). First, participants rated six randomly ordered general affective items, which included *happy*, *threatened*, *frustrated*, *angry*, *sad*, and *proud*. Participants then completed a randomised 6-item scale assessing negative self-directed affect, which included *embarrassment*, *guilt*, *self-critical*, *shameful*, *disappointed with myself*, and *dissatisfied with myself*. Finally, participants completed a randomised 3-item scale assessing negative other-directed affect, which included *angry at others*, *irritated at others*, and *disgusted at others*. The presentation of the negative self-directed affect and the negative other-directed affect measures was randomised.

3.1.2.8 Lexical decision task. The same LDT from Experiment 1 was used in Experiment 2.

3.1.2.9 Instructional manipulation check B. The same instructional manipulation check B from Experiment 1 was used in Experiment 2. In total, 21% of 119 participants failed the instructional manipulation check B (19% failed once, 2% failed twice, and 1% failed six times) before correctly responding and continuing the experiment.

3.1.2.10 Actual–ideal differentiation. The same actual–ideal differentiation measure from Experiment 1 was used in Experiment 2.

3.1.2.11 Goal commitment[9]. To assess goal commitment, participants indicated how committed they were to pursuing the goal to act egalitarian towards Black people using a single item ("How committed are you to pursuing the goal of being egalitarian towards Black people?"), on a 9-point scale ranging from 0 (*not at all committed)* to 8 (*extremely committed).*

3.1.2.12 Goal importance[9]. The measure of goal importance remained as described in Experiment 1 except that the scale was changed to a 9-point scale ranging from 0 (*not at all important)* to 8 (*extremely important)*.

3.1.2.13 Identification with egalitarians towards Black people[10]. This measure was identical to the measure of identification with egalitarian people administered at the beginning of the present experiment, except that this measure specifically assessed participants' identification with people who act egalitarian *towards Black people* (e.g., "I identify strongly with people who are egalitarian towards Black people"). This measure was included to

[9] The goal commitment and goal importance measures were presented along with a measure of self-concordance in a random order. The self-concordance measure was included for exploratory purposes. The method and analyses are presented in Appendix F.

[10] Following the identification with egalitarians towards Black people measure, participants were presented with four exploratory measures to complete: IMS and EMS, contact with Black people, social desirability, and reactance (presented in a random order). The methods and analyses are presented in Appendix F.

explore the effect of reducing the size of the external reference group even further by excluding people who are not egalitarian towards Black people specifically.

3.1.2.14 Positive distinctiveness questions. Finally, the same measure of positive distinctiveness from Experiment 1 was used in Experiment 2, except that the scale was changed to a 9-point scale ranging from 0 (*not at all positive)* to 8 (*extremely positive)*.

After completing these measures, participants provided demographic information and underwent "process" debriefing (McFarland, Cheam, & Buehler, 2007) to mitigate any lasting effects of the negative feedback. First, participants received standard debriefing information (i.e., goals and hypotheses of the experiment). Second, participants were explicitly told that the test and feedback were false, and that the test had been piloted to ensure believability. Third, participants were prompted to avoid finding reasons to support the feedback and to consider how feedback at 100% could equally be true for them. Finally, the experimenter discussed the study with the participants to ensure understanding and to provide the opportunity for questions[11].

3.2 Results and Discussion

3.2.1 Manipulation checks.

3.2.1.1 Goal source primes. A goal prime index was created by averaging responses, with higher scores indicating greater agreement with the goal primes. Participants in the internal goal condition agreed moderately with the goal primes (M = 2.28, SD = 0.62, α = .77); as did participants in the external goal condition (M = 2.35, SD = 0.75, α = .91). Therefore, focusing the external goal primes on not only other egalitarian people's beliefs and

[11] The experimenter did not probe for suspicion in Experiment 1 or in any of the following experiments. However, in a separate experiment on the same topic, participants underwent "funnel" debriefing (Bargh & Chartrand, 2000) to probe suspicion. Using this debriefing, only 3 participants out of 139 participants were actually suspicious of the feedback (61% or 100%) that they received being false.

expectations, but also on egalitarian people's negative reactions to nonegalitarian acts was successful in increasing the internal consistency of the external goal primes and increasing participants' agreement with the external goal primes to the same level as participants' agreement with the internal goal primes.

3.2.1.2 IAT error rates. The percentage of errors that participants made during the critical blocks of the IAT was calculated (M = 4%, SD = 3.40). A one-way between-participants ANOVA indicated that the percentage of errors made during the IAT did not differ between the internal and external goal conditions, $F(1, 117) = 2.28$, $p = .13$, $\eta^2_p = .02$.

3.2.1.3 Goal–behaviour status. Participants' estimates of how well they had performed on the test and how likely they were to achieve the egalitarian goal in the future were analysed to determine how participants interpreted their performance on the test *before* they received the external feedback. Overall, participants indicated that they performed moderately well on the ostensible test (M = 4.62, SD = 1.52). A one-way between-participants ANOVA indicated that participants' perception of how well they had performed on the test did not differ between the internal and external goal conditions, $F(1, 117) = 0.04$, $p = .84$, $\eta^2_p < .001$.

In addition, participants estimated that they were roughly 70% likely to achieve the egalitarian goal (M = 7.50, SD = 1.73; 7.50 was between choice option 61-70% and option 71-80%), matching the subsequent false feedback at 78% fairly well. A one-way between-participants ANOVA indicated that participants' estimates of their likely achievement of the goal did not differ between the internal and external goal conditions, $F(1, 117) = 0.77$, $p = .38$, $\eta^2_p < .01$.

3.2.2 Affect.

3.2.2.1 Negative self-directed affect. An index of negative self-directed affect was created by averaging responses, with higher scores indicating higher negative self-directed affect (M = 1.27, SD = 1.27, α = .91). A one-way between-participants ANOVA indicated that negative self-directed affect did not differ between the internal and external goal conditions, $F(1, 117) = 0.76$, $p = .39$, $\eta^2_p < .01$.

3.2.2.2 Negative other-directed affect. An index of negative other-directed affect was created by averaging responses, with higher scores indicating higher negative other-directed affect (M = 0.61, SD = 1.37, α = .96). A one-way between-participants ANOVA indicated that negative other-directed affect did not differ between the internal and external goal conditions, $F(1, 117) < 0.01$, $p = .97$, $\eta^2_p < .01$.

3.2.2.3 Additional affect analysis. An index of general positive affect was created by averaging participants' responses to *happy* and *proud* (M = 3.79, SD = 1.77; α = .80). A one-way between-participants ANOVA indicated that participants in the internal goal condition (M = 3.43, SE = 0.22) reported significantly lower general positive affect than participants in the external goal condition (M = 4.16, SE = 0.23), $F(1, 117) = 5.36$, $p = .022$, $\eta^2_p = .04$).

An index of general negative affect was created by averaging participants' responses to *threatened, frustrated, angry*, and *sad* (M = 1.11, SD = 1.02, α = .68). A one-way between-participants ANOVA indicated that general negative affect did not differ between the internal and external goal conditions, $F(1, 117) = 0.10$, $p = .75$, $\eta^2_p < .01$.

3.2.3 Stereotype accessibility. The main dependent variable was mean RTs to Black-stereotypic and stereotype-neutral words on the LDT. No participants' data were excluded from the analyses as no participants made incorrect responses on more than 15% of trials. Trials with incorrect responses (4.50% of the data) and reaction times exceeding 2.5 standard

deviations away from each participant's individual mean RT (2.69% of the data) were excluded from the analyses.

The data were submitted to a 2 (Goal Source: internal vs. external) × 2 (Word Type: Black-stereotypic vs. stereotype-neutral) × 2 (Word Valence: negative vs. positive) × 2 (Block: 1 vs. 2) mixed-model ANOVA with goal source as a between-participants factor. The analysis revealed three significant main effects. First, a significant main effect of block, $F(1, 117) = 36.74$, $p < .001$, $\eta^2_p = .24$, indicated that participants responded faster in Block 2 (M = 491 ms, SE = 5.30) than Block 1(M = 516 ms, SE = 6.39). Second, a significant main effect of word valence, $F(1, 117) = 73.82$, $p < .001$, $\eta^2_p = .39$, indicated that participants responded faster to positive words (M = 491 ms, SE = 5.19) than negative words (M = 515 ms, SE = 6.09). Finally, a significant main effect of word type, $F(1, 117) = 20.80$, $p < .001$, $\eta^2_p = .15$, indicated that participants responded faster to Black-stereotypic words (M = 498 ms, SE = 5.29) than stereotype-neutral words (M = 509 ms, SE = 5.92).

The analysis also revealed two significant interactions of theoretical interest. First, there was a significant Word Type × Word Valence interaction, $F(1, 117) = 15.22$, $p < .001$, $\eta^2_p = .12$. For positive words, participants responded equally fast to Black-stereotypic words (M = 491 ms, SE = 5.33) and stereotype-neutral words (M = 492 ms, SE = 5.50), $t(118) = 0.21$, $p = .83$, $d = 0.01$. For negative words, participants responded faster to Black-stereotypic words (M = 504 ms, SE = 5.73) than stereotype-neutral words (M = 525 ms, SE = 6.99), $t(118) = 5.37$, $p < .001$, $d = 0.30$.

Second, and more importantly, the analysis also yielded a significant Goal Source × Word Type interaction, $F(1, 117) = 5.63$, $p = .019$, $\eta^2_p = .05$. Interaction means are presented in Figure 7. Participants primed with an internal goal responded equally fast to Black-stereotypic and stereotype-neutral words, $t(59) = 1.65$, $p = .11$, $d = 0.07$. In contrast,

participants primed with an external goal responded faster to Black-stereotypic words than stereotype-neutral words, $t(58) = 4.15$, $p < .001$, $d = 0.29$. No additional theoretically relevant main effects or interactions were significant[12].

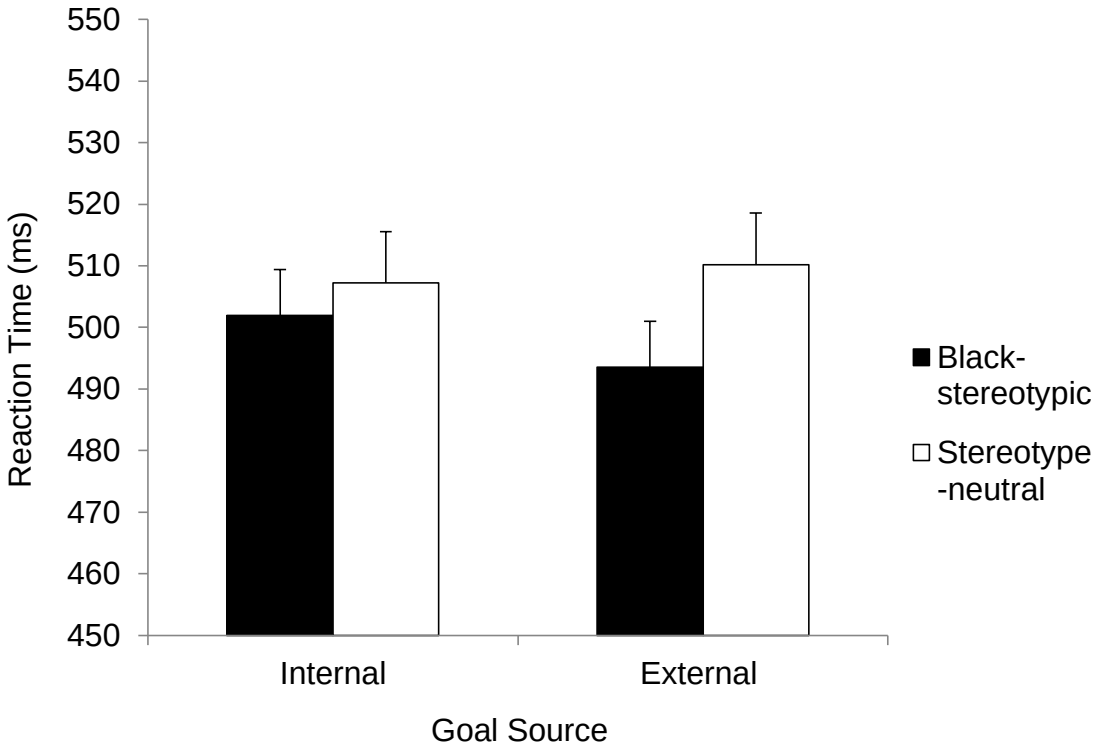

Figure 7. Mean reaction times (ms) as a function of goal source and word type, Experiment 2. Note. Error bars represent standard error.

3.2.4 Potential moderators. A stereotype accessibility index was created to explore the potential moderating role of identification with the external reference group (i.e., egalitarians in the present experiment) and goal importance on the influence of goal source on stereotype accessibility. Specifically, response times to Black-stereotypic words were

[12] Two further significant interactions emerged: a Block × Word Valence interaction, $F(1, 117) = 3.97$, $p = .049$, $\eta^2_p = .03$, and a Block × Word Type × Word Valence interaction, $F(1, 117) = 6.34$, $p = .013$, $\eta^2_p = .05$. As these interactions are not pertinent to the current investigation, further analysis was not conducted. No additional main effects or interactions were significant, $p > .35$.

subtracted from response times to stereotype-neutral words, with higher scores indicating stronger stereotype accessibility (*Grand M* = 44 ms, *SD* = 107.02).

3.2.4.1 Identification. Two indices of identification were created. The first index was created by averaging responses to the identification measure presented at the beginning of the experiment, with higher scores indicating higher identification with egalitarians (*M* = 5.27, *SD* = 1.39, α = .91[13]). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Identification standardised regression analysis. The analysis revealed a significant main effect of goal source, β = -22.59, $t(115)$ = 2.33, p = .022, R^2 = .05; accessibility of Black stereotypes was greater in the external than the internal goal condition. No additional main effects or interactions were significant, all $p > .43$; identification had no effect.

The second index was created by averaging responses to the identification measure presented after the LDT, with higher scores indicating higher identification with people who act egalitarian *towards Black people* (*M* = 5.22, *SD* = 1.83, α = .94[13]). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Identification standardised regression analysis. The analysis revealed a significant main effect of goal source, β = -20.38, $t(115)$ = 2.08, p = .040, R^2 = .06; accessibility of Black stereotypes was greater in the external than the internal goal condition. No additional main effects or interactions were significant, all $p > .16$; identification had no effect.

[13] Consistent with Experiment 1, although one of the items in the identification measures ("I perceive myself as being similar to egalitarian people" in the measure of identification with egalitarians, and "I perceive myself as being similar to people who are egalitarian towards Black people" in the measure assessing identification with egalitarians who act egalitarian towards Black people) could be considered more a measure of prototypicality than of identification, removing these items from the identification indices actually reduced the reliability of the indices (from α = .91 to α = .88, and from α = .94 to α = .93, respectively). Consequently, these items were retained in the identification indices.

3.2.4.2 Goal importance. An index of goal importance was created by averaging responses, with higher scores indicating higher goal importance (M = 6.83, SD = 1.30, α = .93). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Goal Importance standardised regression analysis. The analysis revealed a significant main effect of goal source, β = -23.05, $t(115) = 2.38$, p = .019, R^2 = .05; accessibility of Black stereotypes was greater in the external than the internal goal condition. No additional main effects or interactions were significant, all $p > .38$; goal importance had no effect.

3.2.4.2 Positive distinctiveness. To assess positive distinctiveness as a potential competing goal, the actual–ideal differentiation and the positive distinctiveness data were analysed. For the actual–ideal differentiation data, an actual–ideal differentiation score was calculated by subtracting the ideal score from the actual score, with higher scores indicating a desire for greater differentiation from the Black ethnic group (i.e., greater actual differentiation than ideal differentiation; M = -2.14, SD = 1.43). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Actual–Ideal Differentiation standardised regression analysis. The analysis revealed a significant main effect of goal source, β = -23.20, $t(115) = 2.39$, p = .018, R^2 = .05; accessibility of Black stereotypes was greater in the external than the internal goal condition. No additional main effects or interactions were significant, all $p > .37$; actual–ideal differentiation had no effect.

For the positive distinctiveness data, the two positive distinctiveness items were averaged into an index (M = 4.77, SD = 1.28, α = .64). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Positive Distinctiveness standardised regression analysis. The analysis revealed a significant main

effect of goal source, β = -23.90, $t(115) = 2.39$, $p = .019$, $R^2 = .05$; accessibility of Black stereotypes was greater in the external than the internal goal condition. No additional main effects or interactions were significant, all $p > .68$; positive distinctiveness had no effect.

3.2.5 Additional analyses.

3.2.5.1 Goal commitment. Overall, goal commitment was high ($M = 6.05$, $SD = 1.51$). A one-way between-participants ANOVA indicated that goal commitment did not differ between the internal and external goal conditions, $F(1, 117) = 0.37$, $p = .54$, $\eta^2_p < .01$.

3.2.6 Summary. The findings of Experiment 2 demonstrated that after receiving false feedback that highlighted the potential for a large goal–behaviour discrepancy in the context of egalitarianism, participants primed with internal reasons to act egalitarian exhibited little to no accessibility of Black stereotypes. In contrast, participants primed with external reasons to act egalitarian exhibited accessibility of Black stereotypes. This pattern suggests that participants in the internal goal condition controlled their stereotypes whereas participants in the external goal condition did not. Furthermore, the findings of Experiment 2 demonstrate the opposite pattern to the findings of Experiment 1 (see Table 1). In Experiment 1, after reflecting on a past successful egalitarian act, as opposed to receiving false feedback that highlighted a large goal–behaviour discrepancy in Experiment 2, participants in the internal goal condition exhibited accessibility of Black stereotypes, whereas participants in the external goal condition exhibited little to no accessibility of Black stereotypes.

Table 1
Summary of stereotype accessibility findings as a function of goal source and goal–behaviour status for Experiments 1 & 2

Egalitarian Goal Source	Goal–Behaviour Status	Stereotype Accessibility?	Supported Fishbach and Colleagues' Model?	Experiment #
Internal	Large (78%)	X	✓	Experiment 2
	Small (past success)	✓	✓	Experiment 1
External	Large (78%)	✓	✓	Experiment 2
	Small (past success)	X	✓	Experiment 1

The findings of Experiment 2 are also consistent with my hypotheses, derived from Fishbach and colleagues' model: When participants perceive the potential for a large goal–behaviour discrepancy, participants in the internal goal condition should exhibit increased goal adherence by demonstrating little to no stereotype accessibility. In contrast, participants in the external goal condition should exhibit decreased goal adherence by demonstrating stereotype accessibility.

In relation to affect, Experiment 2 found no evidence of increased negative self-directed affect following the activation of a potential goal–behaviour discrepancy for participants with an internal (vs. external) egalitarian goal. In fact, the level of negative self-directed affect was equally low for participants in the internal and external goal conditions—and yet participants in the internal goal condition still exhibited little to no accessibility of Black stereotypes. This is inconsistent with past research by both Monteith and colleagues (Devine et al., 1991; Monteith, 1993; Monteith et al., 1993; Plant & Devine, 1998; Zuwerink et al., 1996) and Fehr and Sassenberg (2010), which found that both low-prejudiced and internally motivated participants report negative self-directed affect following a large goal–behaviour discrepancy, whereas high-prejudiced and externally motivated participants do not. The finding of the present experiment is particularly inconsistent with Monteith's (1993; Monteith et al., 2002) self-regulation of prejudice model, which states that negative self-

directed affect is a necessary component for people to control their prejudice. This difference may have occurred because Monteith and colleagues' (Devine et al., 1991; Monteith, 1993; Monteith et al., 1993; Plant & Devine, 1998; Zuwerink et al., 1996) measured people's beliefs (i.e., high- or low-prejudiced) and Fehr and Sassenberg (2010), and Plant and Devine (1998), measured people's motivation (i.e., internal or external), which might represent their underlying chronic motivations. In contrast, in the present experiment, internal and external goals were activated temporarily through priming. There may be key differences in the affect that results from a large goal–behaviour discrepancy relating to a chronically accessible goal (as in Monteith & colleagues', Plant & Devine's, and Fehr & Sassenberg's research) versus a temporarily accessible goal (as in the present experiment). For example, negative self-directed affect may be necessary for self-regulation to occur for chronic goals that are measured, but not for temporary goals that are primed.

The present investigation also found no evidence of increased negative other-directed affect following the activation of a potential goal–behaviour discrepancy for participants with an external (vs. internal) egalitarian goal. This finding is contrary to past research demonstrating that high-prejudiced and externally motivated people exhibit negative other-directed affect following a large goal–behaviour discrepancy, whereas low-prejudiced and internally motivated people do not (Devine et al., 1991; Monteith, 1996; Monteith et al., 1993; Plant & Devine, 1998). Researchers have suggested that externally motivated people resent the restriction placed on their autonomy, potentially explaining the occurrence of other-directed affect or angry/threatened affect for externally motivated people in past research (e.g., Brehm, 1966; Legault et al., 2011; Plant & Devine, 1998, 2001). This effect was not replicated within the present investigation, possibly because the participants believed the false feedback they received was based on their own performance from the ostensible egalitarian

test. This may have decreased participants' perception that their autonomy was being restricted in the external goal condition and thus reduced the other-directed affect to the same level as participants in the internal goal condition. This is supported by Fehr and Sassenberg (2010), who informed participants that their performance on a previous task (a modified IAT) indicated that they had more negative attitudes about Arabs than Germans and found no interaction between negative other-directed affect and motivation (i.e., internal or external). Interestingly, in the present experiment, participants in the external goal condition still failed to respond consistently with the egalitarian goal. This suggests that when people with an external egalitarian goal fail to act consistently with the egalitarian goal following a goal–behaviour discrepancy, this is not purely because of the restriction placed on their freedom.

In conclusion, Experiment 2 demonstrated that when people perceive the potential for a large goal–behaviour discrepancy in the context of egalitarianism, only those motivated by internal goals to act egalitarian showed little to no accessibility of Black stereotypes (consistent with stereotype control), even in the absence of negative self-directed affect. In Experiment 3, I examine small goal–behaviour discrepancies in the context of egalitarianism.

4.0 Experiment 3

Experiment 3 aimed to investigate how the source of an egalitarian goal influences stereotype accessibility when there is potentially a small goal–behaviour discrepancy in the context of egalitarianism. Participants were primed with either internal reasons or external reasons to pursue an egalitarian goal. Next participants completed an "egalitarianism test" that ostensibly showed the potential for a small goal–behaviour discrepancy. Afterward, participants completed an LDT to assess stereotype accessibility.

I expected to replicate the findings from Experiment 1 within Experiment 3. Therefore, the predictions in Experiment 1 and 3 are identical. Specifically, Fishbach and

colleagues' model (see Figure 1) proposes that because people with internal goals are committed to their goals, they focus on their goal progress. Perceiving a small goal–behaviour discrepancy should signal that sufficient progress has been made towards the goal, licensing the disengagement from goal pursuit. I, therefore, predicted that when participants in the internal goal condition become aware of the potential for a small goal–behaviour discrepancy, they should exhibit decreased goal adherence by demonstrating stereotype accessibility. Fishbach and colleagues' model also proposes that because people with external goals are uncertainly committed to their goals, they focus on their goal commitment. Perceiving a small goal–behaviour discrepancy should signal that commitment is high, increasing goal adherence. I, therefore, predicted that when participants in the external goal condition become aware of the potential for a small goal–behaviour discrepancy, they should exhibit increased goal adherence by demonstrating little to no stereotype accessibility.

Finally, the same measures of identification with the external reference group, goal importance, goal commitment, and positive distinctiveness that were used in Experiment 2 were also included in Experiment 3. The predictions for identification with the external reference group, goal importance, and goal commitment remain as per Experiments 1 and 2. The predictions for positive distinctiveness were outlined in Experiment 1: Participants in the internal goal condition who also hold a competing positive distinctiveness goal should exhibit greater stereotype accessibility than those who do not hold a competing positive distinctiveness goal because the increased stereotype accessibility would aid the competing positive distinctiveness goal.

I also explored the influence of goal source (i.e., internal or external) on affect following a small goal–behaviour discrepancy using the same measures of affect used in Experiment 2. Based on past research (Devine et al., 1991; Monteith, 1993; Monteith, 1996;

Monteith et al., 1993; Monteith & Voils, 1998; Plant & Devine, 1998; Zuwerink et al., 1993), one might expect to find the following: low negative self-directed affect reported by participants in the internal and external goal conditions; high general positive affect reported by participants in the internal and external goal conditions. One might also expect to find the following: greater other-directed affect reported by participants in the external goal condition than the internal goal condition, and greater general negative affect reported by participants in the external goal condition compared to the internal goal condition, on the basis that the external pressure in the external goal condition may result in reactance due to the restriction of the participants' freedom (Brehm, 1966; Plant and Devine, 1988; 2001).

4.1 Method

4.1.1 Participants and design. Eighty-four White students[14] from the University of Birmingham completed the experiment in exchange for course credit or money (£4). The data from one participant were lost due to a computer malfunction. This left 83 participants (69 female; M_{age} = 19.8 years, SD = 1.84) who were randomly assigned to one of two conditions of a single-factor (Goal Source: internal vs. external) between-participants design.

4.1.2 Materials and procedure. The materials and procedure in Experiment 3 were identical to the materials and procedure used in Experiment 2 except for two changes. First, the categorisation task (the modified IAT; Greenwald et al., 1998) and the false feedback were designed to induce a potentially small goal–behaviour discrepancy in the present experiment, as opposed to a potentially large goal–behaviour discrepancy in Experiment 2.

[14] Based on past research by Fishbach et al. (2011), I expected to achieve a Cohen's *d* somewhere between 0.6 to 0.7 when examining the moderating effect of identification on the influence of goal source on stereotype accessibility. Aprior power analyses (G*Power; Faul et al., 2009) using effect size F^2 (0.09 and 0.13; determined using DeCoster's (2012) effect size conversion calculator), critical alpha (.05), total sample size (119 participants), and number of predictors (2), indicated that 40–60 participants per condition would be required to achieve adequate power at 0.80. In Experiment 3, I collected 40 participants per condition.

Second, the measure of participants' identification with the general category of egalitarian people was moved to later in the session; it replaced the measure of participants' identification with the specific category of people who act egalitarian towards Black people.

4.1.2.1 Small goal–behaviour discrepancy framing task. This version of the modified IAT was identical to the large goal–behaviour discrepancy framing task in Experiment 2 (i.e., with both target blocks composed of prejudice-congruent stimulus combinations and error feedback), but the task instructions were altered. In the small goal–behaviour discrepancy framing task, participants learned that they would be completing a test of how motivated they were to act egalitarian and would receive feedback about their performance, based on how *accurately but not how quickly* they responded. Lifting the time constraint was meant to result in fewer errors[15] and heighten the perception of successful performance on the test.

4.1.2.1 98% feedback. Participants were presented with a graph depicting their performance; an arrow extended upwards to 98% on a scale from 0–100% (see Figure 8). The accompanying wording was consistent with each goal source.

[15] This assertion is based on research by Fiedler and Bluemke (2005; Experiment 1). German participants completed a typical German–Turkish IAT twice. The second time, they were explicitly instructed to slow down on both prejudice-congruent and -incongruent trials to avoid appearing prejudiced towards Turkish people. Analysis of the error rates indicated that the number of errors decreased after participants received instructions to slow down prior to repeating the IAT.

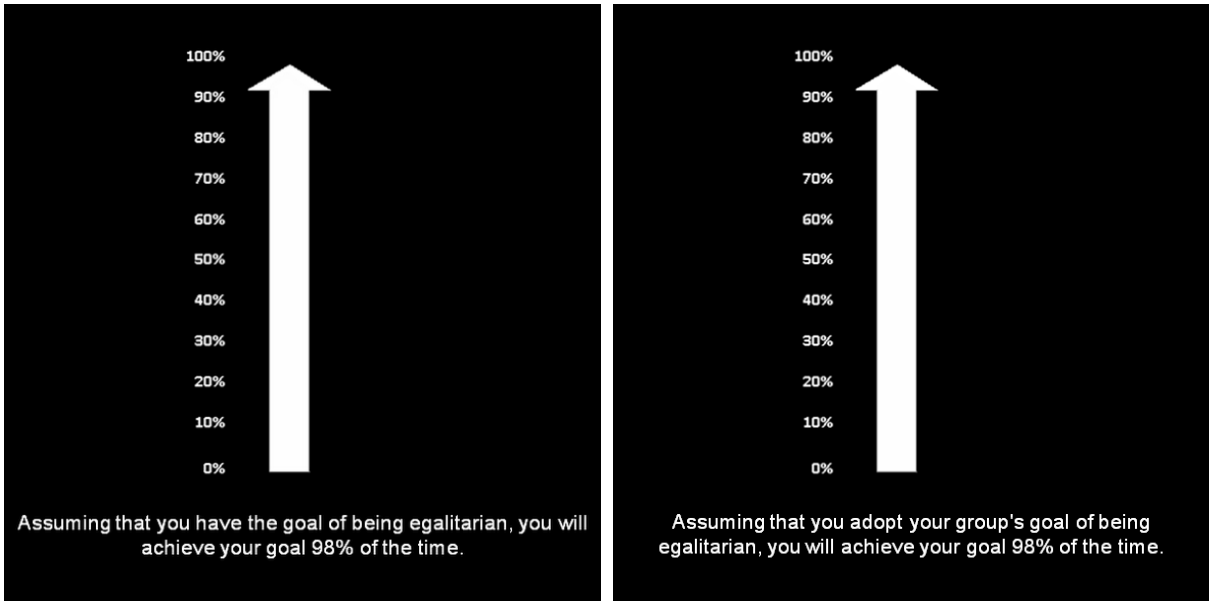

Figure 8. **98% feedback that participants received to induce a potentially small goal–behaviour discrepancy, Experiment 3.** The left panel depicts the feedback in the internal goal condition while the right depicts the external goal condition.

4.2 Results and Discussion

4.2.1 Manipulation checks.

4.2.1.1 Instructional manipulation checks. For instructional manipulation check A, a total of 32% failed (16% failed once, 8% failed twice, 4% failed three times, 4% failed five times, and 1% failed six times). For instructional manipulation check B, a total of 13% failed (11% failed once, 1% failed twice, and 1% failed six times).

4.2.1.2 Goal source primes. A goal prime index was created by averaging responses, with higher scores indicating greater agreement with the goal primes. Participants in the internal goal condition agreed moderately with the internal goal primes (M = 2.29, SD = 0.52, α = .61[16]); as did participants in the external goal condition (M = 2.28, SD = 0.69, α = .82).

[16] I believe that the lower internal consistency in Experiment 3 (α = .61) compared to Experiments 1 and 2 (α = .87 and α = .77, respectively) is due to some participants missing the reverse coding of three of the internal goal primes ("According to my personal values, being intolerant of Black people is ok", "It is my personal belief that Black people should NOT have an equal chance or an equal say", and "I believe people should NOT be concerned

4.2.1.3 IAT error rates. The percentage of errors that participants made during the critical blocks of the IAT was calculated (M = 1%, SD = 1.42). A one-way between-participants ANOVA indicated that the percentage of errors made during the IAT did not differ between the internal and external goal conditions, $F(1, 81) = .28$, $p = .60$, $\eta^2_p < .01$. An independent samples t-test indicated that participants made significantly fewer errors during the IAT in Experiment 3 (designed to induce a small goal–behaviour discrepancy; M = 1%, SE = 0.16) compared to Experiment 2 (designed to induce a large goal–behaviour discrepancy; M = 4%, SE = 0.31), $t(169.10) = 7.64$, $p < .001$, $d = 1.02$[17]. This finding indicates that the change made to the instructions for the egalitarian "test" in Experiment 3 successfully reduced errors.

4.2.1.4 Goal–behaviour status. Participants' estimates of how well they had performed on the test and how likely they were to achieve the egalitarian goal in the future were analysed to determine how participants interpreted their performance on the test *before* receiving the external feedback. Overall, participants indicated that they performed moderately to very well on the ostensible test (M = 5.36, SD = 1.45). A one-way between-participants ANOVA indicated that participants' perception of how well they had performed on the test did not differ between the internal and external goal conditions, $F(1, 81) = 1.54$, $p = .22$, $\eta^2_p = .02$.

In addition, participants estimated that they were 75% likely to achieve the egalitarian goal (M = 8.06, SD = 1.70; 8.06 was choice option 71-80%). A one-way between-participants

---

about the welfare of Black people"). Indeed, examining the *Cronbach's Alpha if Item Deleted* column indicated that removing these reversed internal goal primes would have increased the cronbach's alpha ($\alpha = .65$, $\alpha = .62$, and $\alpha = .63$, respectively), whereas removing any of the other internal goal primes would have decreased the cronbach's alpha. In contrast, removing these reversed internal goal primes in Experiments 1 or 2 would not have increased the cronbach's alpha.

[17] Levene's Test for Equality of Variances was significant ($F = 38.09$, $p < .001$); therefore, a t-test not assuming equal variances was conducted.

ANOVA indicated that participants' estimates of their likely achievement of the goal did not differ between the internal and external goal conditions, $F(1, 81) = 1.22$, $p = .27$, $\eta^2_p = .02$.

Further, participants' estimates of their likely achievement of the goal were similar in Experiment 2 and 3 (roughly 70% vs. 75%, respectively). Although an independent sample *t*-test indicated that participants in Experiment 3 estimated their likely achievement of the goal significantly higher than participants in Experiment 2, $t(200) = 2.27$, $p = .024$, $d = 0.32$, the difference between participants' estimates in Experiment 2 and 3 was not as large as I anticipated. Importantly, participants received false feedback *after* estimating their likely achievement of the goal. Therefore, receiving feedback at 78% (Experiment 3) *after* estimating their likely achievement of the goal at 70% (on average) should have matched their discrepancy expectations and resulted in participants perceiving a goal–behaviour discrepancy. In contrast, receiving feedback at 98% (Experiment 2) *after* estimating their likely achievement of the goal at 75% (on average) should have *exceeded* their expectations and resulted in participants perceiving a smaller goal–behaviour discrepancy than in Experiment 2.

4.2.2 Affect.

4.2.2.1 Negative self-directed affect. An index of negative self-directed affect was created by averaging participants responses to five of the six affect items assessing negative self-directed affect (*embarrassment*, *guilt*, *shameful*, *disappointed with myself*, and *dissatisfied with myself)*[18], with higher scores indicating higher negative self-directed affect ($M = 0.22$, $SD = 0.38$, $\alpha = .74$). A one-way between-participants ANOVA indicated that

[18] When averaging participants responses to all six of the affect items assessing negative self-directed affect (*embarrassment, guilt, self-critical, shameful, disappointed with myself,* and *dissatisfied with myself),* internal consistency was low ($M = 0.41$, $SD = 0.48$, $\alpha = .52$).

negative self-directed affect did not differ between the internal and external goal conditions, $F(1, 81) = 0.08$, $p = .78$, $\eta^2_p < .01$.

4.2.2.2 Negative other-directed affect. An index of negative other-directed affect was created by averaging responses, with higher scores indicating higher negative other-directed affect ($M = 0.46$, $SD = 1.22$, $\alpha = .95$). A one-way between-participants ANOVA indicated that negative other-directed affect did not differ between the internal and external goal conditions, $F(1, 81) < .001$, $p = .99$, $\eta^2_p < .001$.

4.2.2.3 Additional affect analysis. An index of general positive affect was created by averaging participants' responses to *happy* and *proud* ($M = 4.85$, $SD = 1.91$; $\alpha = .78$). A one-way between-participants ANOVA indicated that general positive affect did not differ between the internal and external goal conditions, $F(1, 81) = 0.13$, $p = .72$, $\eta^2_p < .01$.

An index of general negative affect was created by averaging participants' responses to *threatened*, *frustrated*, *angry*, and *sad* ($M = 0.41$, $SD = 0.86$, $\alpha = .72$). A one-way between-participants ANOVA indicated that general negative affect did not differ between the internal and external goal conditions, $F(1, 81) = 0.28$, $p = .87$, $\eta^2_p < .001$.

4.2.3 Stereotype accessibility. The main dependent variable was mean RTs to Black-stereotypic and stereotype-neutral words on the LDT. One participant's data were excluded from the analyses as this participant made incorrect responses on more than 15% of trials. For the remaining 82 participants, trials with incorrect responses (3.72% of the data) and reaction times exceeding 2.5 standard deviations away from each individual participant's mean RT (2.90% of the data) were excluded from the analyses.

The data were submitted to a 2 (Goal Source: internal vs. external) × 2 (Word Type: Black-stereotypic vs. stereotype-neutral) × 2 (Word Valence: negative vs. positive) × 2 (Block: 1 vs. 2) mixed-model ANOVA with goal source as a between-participants factor. The

analysis revealed three significant main effects. First, a significant main effect of block, $F(1, 80) = 20.84$, $p < .001$, $\eta^2_p = .21$, indicated that participants responded faster in Block 2 (M = 521 ms, SE = 7.91) than Block 1 (M = 548 ms, SE = 9.31). Second, a significant main effect of word valence, $F(1, 80) = 61.46$, $p < .001$, $\eta^2_p = .43$, indicated that participants responded faster to positive words (M = 521 ms, SE = 8.38) than negative words (M = 548 ms, SE = 8.25). Finally, a significant main effect of word type, $F(1, 80) = 4.23$, $p = .043$, $\eta^2_p = .05$, indicated that participants responded faster to Black-stereotypic words (M = 531 ms, SE = 8.31) than stereotype-neutral words (M = 538 ms, SE = 8.29).

The analysis also revealed a significant Word Type × Word Valence interaction, $F(1, 80) = 14.14$, $p < .001$, $\eta^2_p = .15$. Interaction means are presented in Figure 9. For positive words, participants responded marginally faster to stereotype-neutral than Black-stereotypic words, $t(81) = 1.94$, $p = .056$, $d = .10$. For negative words, participants responded faster to Black-stereotypic than stereotype-neutral words, $t(81) = 3.56$, $p < .001$, $d = .26$. However, the theoretically important main effect of goal source, $F(1, 80) = 0.04$, $p = .85$, $\eta^2_p < .001$, and the interactions between Goal Source × Word Type, $F(1, 80) = 0.06$, $p = .82$, $\eta^2_p < .01$, and Goal Source × Word Type × Word Valence, $F(1, 80) = 0.37$, $p = .54$, $\eta^2_p < .01$, were all non-significant[19]; goal source had no main or interaction effects.

[19] There was also a significant Block × Word Valence interaction, $F(1, 80) = 5.68$, $p = .020$ $\eta^2_p = .07$. As this interaction is not pertinent to the current investigation, further analysis was not conducted. No other main effects or interactions were significant, $p > .10$.

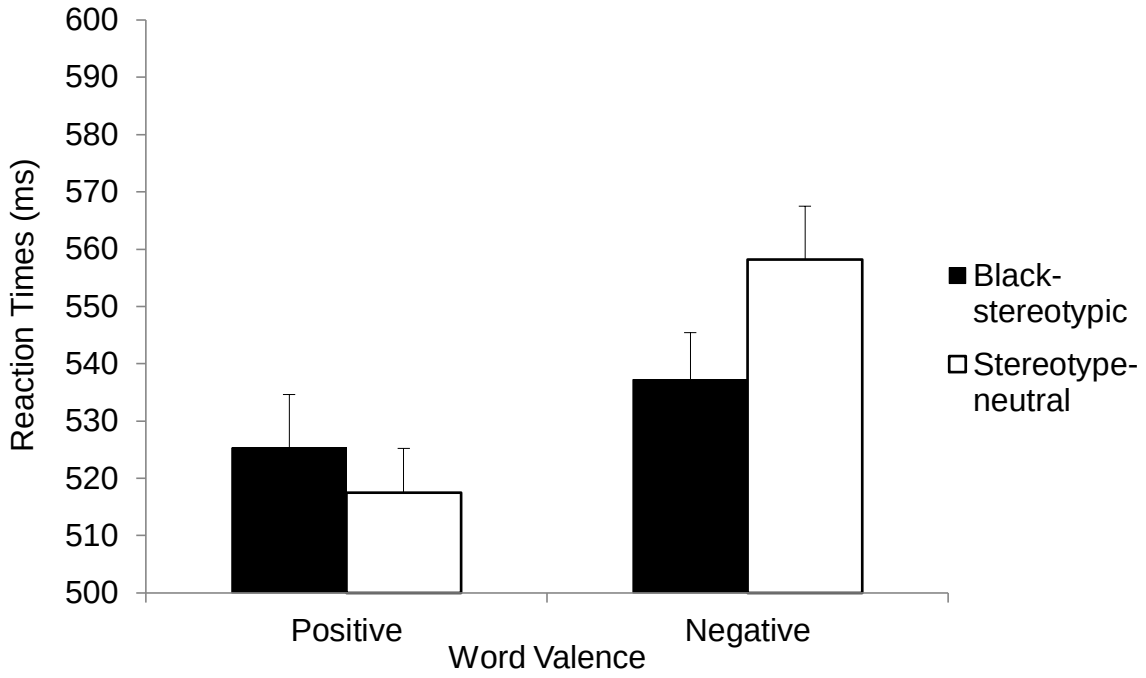

Figure 9. **Mean reaction times (ms) as a function of word valence and word type, Experiment 3.** Note. Error bars represent standard error.

4.2.4 Potential moderators. A negative stereotype accessibility index was created as in Experiment 2. Specifically, response times to negative Black-stereotypic words were subtracted from response times to negative stereotype-neutral words, with higher scores indicating stronger negative stereotype accessibility (*Grand M* = 42 ms, *SD* = 105.07).

4.2.4.1 Identification with egalitarians. An index of identification with egalitarians was created by averaging responses, with higher scores indicating higher identification with egalitarians (*M* = 5.57, *SD* = 1.73, α = .95[20]). The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) ×

[20] Consistent with Experiments 1 and 2, although one of the items in the identification measure ("I perceive myself as being similar to egalitarian people") could be considered more a measure of prototypicality than of identification, removing this item from the identification index actually reduced the reliability of the index (from α = .86 to α = .80). Consequently, this item was retained in the identification index.

Identification standardised regression analysis. No main effects or interactions were significant, all $p > .85$; identification had no effect.

4.2.4.2 Goal importance. An index of goal importance was created by averaging responses, with higher scores indicating higher goal importance ($M = 6.86$, $SD = 1.27$, $\alpha = .89$). The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Goal Importance standardised regression analysis. No main effects or interactions were significant, all $p > .11$; goal importance had no effect.

4.2.4.3 Positive distinctiveness. To assess positive distinctiveness as a potential competing goal, the actual–ideal differentiation and the positive distinctiveness data were analysed. For the actual–ideal data, an actual–ideal differentiation score was calculated by subtracting the ideal score from the actual score, with higher scores indicating a desire for greater differentiation from the Black ethnic group (i.e., greater actual differentiation than ideal differentiation; $M = 1.86$, $SD = 1.31$). The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Actual–Ideal Differentiation standardised regression analysis. No main effects or interactions were significant, all $p > .64$; actual–ideal differentiation had no effect.

For the positive distinctiveness data, the two items assessing positive distinctiveness were not indexed due to poor internal consistency ($\alpha = .32$). Consequently, two standardised regression analyses (Goal Source × Item) were conducted to examine each positive distinctiveness item separately. In neither analysis were any main effects or interactions significant, all $p > .29$; positive distinctiveness had no effect.

4.2.5 Additional analyses.

4.2.5.1 Goal commitment. Overall, goal commitment was high (M = 6.06, SD = 1.62). A one-way between-participants ANOVA indicated that goal commitment did not differ between the internal and external goal conditions, $F(1, 81) = 0.22$, $p = .64$, $\eta^2_p < .01$.

4.2.6 Summary. The findings of Experiment 3 demonstrated that after receiving false feedback that induced a potentially small goal–behaviour discrepancy in the context of egalitarianism, participants in both the internal and external goal conditions exhibited accessibility of negative Black stereotypes. This pattern suggests that neither participants in the internal nor the external goal conditions controlled their stereotypes. In contrast, in Experiment 2, in which participants received false feedback that highlighted a large goal–behaviour discrepancy, participants in the internal goal condition exhibited little to no accessibility of Black stereotypes, whereas participants in the external goal condition exhibited accessibility of Black stereotypes (see Table 2). Furthermore, the findings of Experiment 3 do not replicate the findings of Experiment 1 as expected. In Experiment 1, after reflecting on a past successful egalitarian act, as opposed to receiving false feedback that highlighted a small goal–behaviour discrepancy in Experiment 3, participants in the internal goal condition exhibited accessibility of Black stereotypes, whereas participants in the external goal condition exhibited little to no accessibility of Black stereotypes.

Table 2
Summary of stereotype accessibility findings as a function of goal source and goal–behaviour status for Experiments 1–3.

Egalitarian Goal Source	Goal–Behaviour Status	Stereotype Accessibility?	Supported Fishbach and Colleagues' Model?	Experiment #
Internal	Large (78%)	X	✓	Experiment 2
	Small (past success)	✓	✓	Experiment 1
	Small (98%)	✓	✓	Experiment 3
External	Large (78%)	✓	✓	Experiment 2
	Small (past success)	X	✓	Experiment 1
	Small (98%)	✓	X	Experiment 3

The stereotype accessibility for participants in the internal goal condition supports my hypothesis, derived from Fishbach and colleagues' model: When participants perceive the potential for a small goal–behaviour discrepancy, participants in the internal goal condition exhibited decreased goal adherence by demonstrating stereotype accessibility. However, the stereotype accessibility for participants in the external goal condition contradicts my hypothesis, also derived from Fishbach and colleagues' model: When participants perceive the potential for a small goal–behaviour discrepancy, participants in the external goal should exhibit increased goal adherence by demonstrating little to no stereotype accessibility.

There are two possible explanations for why participants in the external goal condition exhibited stereotype activation rather than control. First, Moskowitz and Li's (2011; see also Moskowitz, 2002) theorising may be true. Contemplating successful egalitarian acts may affirm one's identity as an egalitarian person, leading to semantic activation of the concept of "egalitarianism" and the relaxation of the egalitarian goal (i.e., efforts to control stereotype activation cease). In the present experiment, the accessibility of negative Black stereotypes for participants in both the internal and external goal condition after perceiving the potential for a small goal–behaviour discrepancy is consistent with Moskowitz and Li's theory. That being

said, the findings of Experiment 1 in the present thesis did not support Moskowitz and Li's theorising. In my Experiment 1, participants primed with external reasons for acting egalitarian exhibited little to no accessibility of Black stereotypes after reflecting on a past successful egalitarian act.

Another possibility is that the percentage required on the egalitarian "test" to induce a sense of progress or success might differ for people with internal and external egalitarian goals. People with an internal egalitarian goal may accumulate benefits as they progress towards the goal (i.e., cumulative goal; see Garvey, 2011). Specifically, each time a person acts consistently with their internal egalitarian goal, they affirm their identity as an egalitarian person and alleviate the tension associated with failing to pursue the goal, even if goal attainment has not yet been achieved (see Moskowitz & Li, 2011; Moskowitz et al., 2011). In contrast, due to social pressure to act egalitarian, people with an external egalitarian goal may benefit only when goal attainment is achieved (e.g., all-or-nothing goal; see Garvey 2011). Specifically, being 98% likely to achieve an external egalitarian goal may not be enough to satisfy social pressure to conform to the egalitarian norm. A goal–behaviour discrepancy of just 2% still indicates that, at times, the person will fail to respond consistently with their egalitarian goal and the person may still experience social sanctions associated with failing to comply with externally imposed egalitarian standards. Indeed, Soman and Cheema (2004; see also Garvey, 2011) argue that all-or-nothing goals are binary in nature, in that people either successfully achieve or fail to achieve the goal. For an external egalitarian goal then, a goal–behaviour discrepancy, regardless of the magnitude, may prevent goal attainment and is likely be perceived as a failure. Therefore, Experiment 4 aimed to examine whether increasing the false feedback to 100% success (i.e., so that there is no discrepancy between participants'

egalitarian behaviour and the external egalitarian standard) results in less stereotype accessibility for people with an external egalitarian goal.

5.0 Experiment 4

Experiment 4 aimed to investigate how the source of an egalitarian goal influences stereotype accessibility when there is potentially a match (i.e., no discrepancy) between participants' actual behaviour and their standards for behaviour in the context of egalitarianism. Experiment 4 was the most conceptually similar to Experiment 1 where participants contemplated a past goal success, only here the goal–behaviour match was highlighted by external feedback and was ostensibly relevant to future behaviour. Specifically, participants were primed with either internal or external reasons to pursue an egalitarian goal. Next participants completed an "egalitarianism test" that ostensibly highlighted the potential for a goal–behaviour match. Afterward, participants completed an LDT to assess stereotype accessibility.

Although Fishbach and colleagues' model only makes predictions about small and large goal–behaviour discrepancies, I reasoned that they might predict the same effects for small goal–behaviour discrepancies and goal–behaviour matches. Fishbach and colleagues' model (see Figure 1) proposes that because people with internal goals are committed to their goals, they focus on their goal progress. Perceiving a goal–behaviour match should signal that sufficient progress has been made towards the goal, licensing the disengagement from goal pursuit. Based on Fishbach and colleagues' reasoning, when participants in the internal goal condition become aware of the potential for a goal–behaviour match, they should exhibit decreased goal adherence by demonstrating stereotype accessibility. Fishbach and colleagues' model also proposes that because people with external goals are uncertainly committed to their goals, they focus on their goal commitment. Perceiving a goal–behaviour match should

signal that commitment is high, increasing goal adherence. Based on Fishbach and colleagues' reasoning, when participants in the external goal condition become aware of the potential for a goal–behaviour match, they should exhibit increased goal adherence by demonstrating little to no stereotype accessibility.

However, the findings of Experiment 3 indicated that Moskowitz and Li's (2011; see also Moskowitz, 2002) theorising, that contemplating success leads to the semantic activation of the concept of "egalitarianism" and the relaxation of the egalitarian goal regardless of goal source, still remains plausible. Receiving external feedback that makes participants feel successful, whether that is 98% or 100%, may result in decreased goal adherence (i.e., stereotype accessibility on an LDT) for people with either internal or external egalitarian goals. Therefore, I made no explicit predictions about how perceiving the potential for a goal–behaviour match would affect stereotype accessibility for people with internal or external egalitarian goals.

Finally, the same measures of identification with the external reference group, goal importance, goal commitment, positive distinctiveness, and affect that were used in Experiments 2 and 3 were used in Experiment 4. The predictions for identification with the external reference group, goal importance, and goal commitment remain as per Experiments 1–3. The predictions for positive distinctiveness and affect remain as per Experiment 3.

5.1 Method

5.1.1 Participants and design. Fifty-three students[21] from the University of Birmingham and 15 students from Cadbury Sixth Form College in Birmingham[22] completed

[21] In Experiment 5, I collected only 25 participants per condition as I focused on the main stereotype accessibility analyses rather than the moderator analyses which have not revealed any significant findings up to this point.

[22] After removing non-White participants, only 6 students from the Cadbury Sixth Form College remained in the data set. Consequently, there were not enough students from the

the experiment in exchange for course credit or money (£6). Ten participants who indicated that they were an ethnicity other than White were removed from the data analyses[23]. This left 58 White students (42 female; M_{age} = 21.1 years, *SD* = 3.82). Participants were randomly assigned to one of two conditions of a single-factor (Goal Source: internal vs. external) between-participants design.

5.1.2 Materials and procedure. The materials and procedure in Experiment 4 were identical to the materials and procedure in Experiment 3 except for three changes. First, the categorisation task (the modified IAT; Greenwald et al., 1998) designed to induce a potentially small goal–behaviour discrepancy in Experiment 2 was improved and the false feedback was increased to 100%. Second, a question was added after participants received the false feedback to gauge whether participants interpreted the feedback as a success or a failure. Third, some minor improvements to the wording of the positive distinctiveness measure were made.

5.1.2.1 Goal–behaviour match framing task. In this version of the modified IAT, participants learned that they would be completing a test of how motivated they were to act egalitarian and would receive feedback about their performance, based on how *accurately but not how quickly* they responded. Instead of completing two critical blocks of prejudice-*congruent* pairings as in Experiments 2 and 3, participants completed two critical blocks of only prejudice-*incongruent* pairings to heighten participants' awareness of their own potential

---

Cadbury Sixth Form College left to examine the two samples (i.e., University of Birmingham and Cadbury Sixth Form College) separately. Adding the 6 students from Cadbury Sixth Form College into the data analyses did not change the overall findings of this experiment; therefore, I collapsed across the two samples.

[23] Of the 10 participants who self-reported an ethnicity other than White, two were Middle-Eastern Asian, five were Asian, one was Black, and two were of mixed ethnicity. Removing these participants did not change the overall findings of this experiment. As the stimuli used in the IAT, and the questionnaire measures (e.g., actual-ideal differentiation measure), drew direct comparisons between the White and Black ethnic group, only White participants were retained in the analyses.

for egalitarian responding. In addition, participants received correct-response feedback in the form of a green tick each time they responded correctly, to increase their awareness of their correct answers during the test. In addition, to increase the believability of the subsequent feedback at 100%, the supposed calculation performed by the computer was also explained. Participants read:

> *After the test, the computer will calculate your result, based on how accurately you respond during the test, not how quickly. Based on past research, we can now determine the percentage of errors that are accidental and can adjust the result so only your deliberate errors are taken into account. This means the feedback you receive will demonstrate a truer reflection of how motivated you are to act egalitarian (i.e., acting fair, just and tolerant) towards Black people*.

5.1.2.2 100% feedback. Participants were presented with a graph depicting their performance; an arrow extended upwards to 100% on a scale from 0–100% (see Figure 10). The accompanying wording was consistent with each goal source.

Figure 10. **100% feedback that participants received to highlight the potential for a goal–behaviour match, Experiment 4.** The left panel depicts the feedback in the internal goal condition while the right depicts the external goal condition.

5.1.2.3 Interpretation of feedback. Immediately after receiving the false feedback, participants indicated how well they had performed on the test, on a 7-point scale ranging from -3 (*completely failed*) to +3 (*completely succeeded*).

5.1.2.4 Positive distinctiveness questions. The first positive distinctiveness question was reworded to ensure understanding. The original wording was "How positive is your ethnic group relative to the Black ethnic group, as viewed by other members of your ethnic group" and required perspective taking. The question was reworded to "How positive would your ethnic group say they are in comparison to the Black group". The second positive distinctiveness question remained unchanged (i.e., "How positively is your group viewed?").

5.2 Results and Discussion

5.2.1 Manipulation checks.

5.2.1.1 Instructional manipulation checks. For instructional manipulation check A, a total of 40% failed (22% failed once, 10% failed twice, 2% failed three, four, five, and six

times). For instructional manipulation check B, a total of 19% failed (17% failed once, and 2% failed twice).

5.2.1.2 Goal source primes. A goal prime index was created by averaging responses, with higher scores indicating greater agreement with the goal primes. Participants in the internal goal condition agreed moderately with the goal primes (M = 2.10, SD = 0.59, α = .63[24]); as did participants in the external goal condition (M = 2.22, SD = 0.77, α = .88).

5.2.1.3 IAT error rates. The percentage of errors that participants made during the critical blocks of the IAT was calculated (M = 2%, SD = 1.56). A one-way between-participants ANOVA indicated that the percentage of errors made during the IAT did not differ between the internal and external goal conditions, $F(1, 56) = 2.18$, $p = .15$, $\eta^2_p = .04$. An independent samples t-test indicated that participants made significantly fewer errors during the IAT in Experiment 4 (designed to induce a goal–behaviour match; M = 2%, SE = 0.32) compared to Experiment 2 (designed to induce a large goal–behaviour discrepancy; M = 4%, SE = 0.31), $t(169.34) = 3.86$, $p < .001$, $d = 0.57$[25]. This finding indicates that the changes made to the egalitarian "test" in Experiment 4 (i.e., only prejudice-incongruent trials, and correct-response feedback) successfully reduced errors.

[24] I believe that the lower internal consistency in Experiment 4 (α = .63) compared to Experiments 1 and 2 (α = .87 and α = .77, respectively) is due to some participants missing the reverse coding of one of the internal goal primes ("According to my personal values, being intolerant of Black people is ok"). Indeed, examining the *Cronbach's Alpha if Item Deleted* column indicated that removing this reversed internal goal prime would have increased the cronbach's alpha (α = .70), whereas removing any of the other internal goal primes would have had no effect or decreased the cronbach's alpha. In contrast, removing this reversed internal goal prime in Experiments 1 or 2 would not have increased the cronbach's alpha.

[25] Levene's Test for Equality of Variances was significant (F = 3.82, p = .052); therefore, a t-test not assuming equal variances was conducted.

5.2.1.4 Goal–behaviour status.

5.2.1.4.1 Estimates before feedback. Participants' estimates of how well they had performed on the test and how likely they were to achieve the egalitarian goal in the future were analysed to determine how participants interpreted their performance on the test *before* receiving the external feedback. Overall, participants indicated that they performed very well on the ostensible test (M = 5.78, SD = 1.48). A one-way between-participants ANOVA indicated that participants perception of how well they had performed on the test did not differ between the internal and external goal conditions, $F(1, 56) = 0.02$, $p = .899$, $\eta^2_p < .001$.

In addition, participants estimated that they were roughly 80% likely to achieve the egalitarian goal (M = 8.52, SD = 1.47; 8.52 was between choice option 71–80% and option 81-90%). A one-way between-participants ANOVA indicated that participants' estimates of their likely achievement of the goal did not differ between the internal and external goal conditions, $F(1, 56) = 0.39$, $p = .54$, $\eta^2_p < .01$.

5.2.1.4.2 Interpretation of the feedback. More importantly, participants' interpretation of the feedback was examined to determine more accurately the size of the goal–behaviour discrepancy induced by the feedback. Participants interpreted the 100% feedback as indicating that they had succeeded completely on the test (M = 2.69, SD = 0.71). A one-way between-participants ANOVA indicated that participants interpretation of the 100% feedback did not differ between the internal and external goal conditions, $F(1, 56) = 0.39$, $p = .53$, $\eta^2_p < .01$.

5.2.2 Affect.

5.2.2.1 Negative self-directed affect. An index of negative self-directed affect was created by averaging responses, with higher scores indicating higher negative self-directed affect (M = 0.32, SD = 0.60, $\alpha = .79$). A one-way between-participants ANOVA indicated that

negative self-directed affect did not differ between the internal and external goal conditions, $F(1, 56) < .01$, $p = .97$, $\eta^2_p < .001$.

5.2.2.2 Negative other-directed affect. An index of negative other-directed affect was created by averaging responses, with higher scores indicating higher negative other-directed affect ($M = 0.41$, $SD = 1.15$, $\alpha = .97$). A one-way between-participants ANOVA indicated that negative other-directed affect did not differ between the internal and external goal conditions, $F(1, 56) = 0.08$, $p = .78$, $\eta^2_p < .01$.

5.2.2.3 Additional affect analysis. An index of general positive affect was created by averaging participants' responses to *happy* and *proud* ($M = 4.63$, $SD = 2.01$; $\alpha = .82$). A one-way between-participants ANOVA indicated that general positive affect did not differ between the internal and external goal conditions, $F(1, 56) = 3.18$, $p = .080$, $\eta^2_p = .05$.

An index of general negative affect was created by averaging participants' responses to *threatened*, *frustrated*, *angry*, and *sad* ($M = 0.39$, $SD = 1.00$, $\alpha = .89$). A one-way between-participants ANOVA indicated that general negative affect did not differ between the internal and external goal conditions, $F(1, 56) = 0.48$, $p = .49$, $\eta^2_p < .01$.

5.2.3 Stereotype accessibility. The main dependent variable was mean RTs to Black-stereotypic and stereotype-neutral words on the LDT. One participants' data were excluded from the analyses as this participant made incorrect responses on more than 15% of trials. For the remaining 57 participants, trials with incorrect responses (5.07% of the data) and reaction times exceeding 2.5 standard deviations away each individual participant's mean RT (2.54% of the data) were excluded from the analyses.

The data were submitted to a 2 (Goal Source: internal vs. external) × 2 (Word Type: Black-stereotypic vs. stereotype-neutral) × 2 (Word Valence: negative vs. positive) × 2 (Block: 1 vs. 2) mixed-model ANOVA with goal source as a between-participants factor. The

analysis revealed two significant main effects. First, a significant main effect of block, $F(1, 55) = 40.07$, $p < .001$, $\eta^2_p = .42$, indicated that participants responded faster in Block 2 (M = 489 ms, SE = 7.24) than Block 1(M = 524 ms, SE = 9.17). Second, a significant main effect of word valence, $F(1, 55) = 35.92$, $p < .001$, $\eta^2_p = .40$, indicated that participants responded faster to positive words (M = 497 ms, SE = 7.44) than negative words (M = 517 ms, SE = 8.48).

The analysis also revealed a significant Word Type × Word Valence interaction, $F(1, 55) = 9.87$, $p = .003$, $\eta^2_p = .15$. Interaction means are presented in Figure 11. For positive words, participants responded faster to stereotype-neutral words than Black-stereotypic words, $t(56) = 2.81$, $p = .007$, $d = .19$. For negative words, participants responded faster to Black-stereotypic than stereotype-neutral words, $t(56) = 2.09$, $p = .041$, $d = .16$. However, the theoretically important main effect of goal source, $F(1, 55) = 1.15$, $p = .29$, $\eta^2_p = .02$, and the interactions between Goal Source × Word Type, $F(1, 55) = 0.18$, $p = .68$, $\eta^2_p < .01$, and Goal Source × Word Type × Word Valence, $F(1, 55) = 0.01$, $p = .92$, $\eta^2_p < .01$, were all non-significant; goal source had no main or interaction effects.

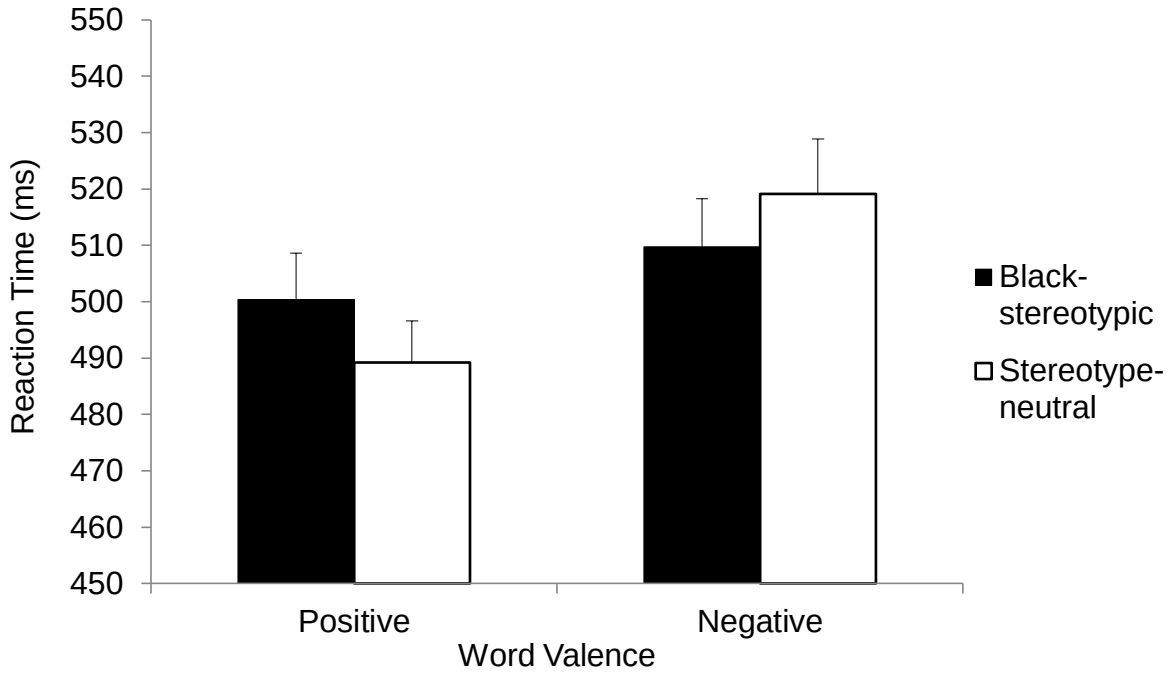

Figure 11. **Mean reaction times (ms) as a function of word valence and word type, Experiment 4.** *Note*. Error bars represent standard error.

5.2.4 Potential moderators. A negative stereotype accessibility index was created as in Experiments 2 and 3. Specifically, response times to negative Black-stereotypic words were subtracted from response times to negative stereotype-neutral words, with higher scores indicating stronger negative stereotype accessibility (*Grand M* = 20 ms, *SD* = 73.63).

5.2.4.1 Identification with egalitarians. An index of identification with egalitarians was created by averaging responses, with higher scores indicating higher identification with egalitarians (*M* = 5.14, *SD* = 1.79, α = .94[26]). The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) ×

[26] Consistent with Experiments 1–3, although one of the items in the identification measure ("I perceive myself as being similar to egalitarian people") could be considered more a measure of prototypicality than of identification, removing this item from the identification index actually reduced the reliability of the index slightly (from α = .94 to α = .93). Consequently, this item was retained in the identification index.

Identification standardised regression analysis. No main effects or interactions were significant, all $p > .069$; identification had no effect.

5.2.4.2 Goal importance. An index of goal importance was created by averaging responses, with higher scores indicating higher goal importance ($M = 6.55$, $SD = 1.54$, $\alpha = .85$). The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Goal Importance standardised regression analysis. There was a significant Goal Source × Goal Importance interaction, $\beta = 20.14$, $t(53) = 2.03$, $p = .047$, $R^2 = .07$. Interaction means are presented in Figure 12. However, when the interaction was decomposed by examining goal importance one standard deviation above and below the mean, neither simple slope was significant. For participants high in goal importance, the goal source priming manipulation had no effect on stereotype accessibility, $\beta = 18.21$, $t(53) = 1.32$, $p = .19$, $R^2 = .07$. For participants low in goal importance, the goal source priming manipulation had no effect on stereotype accessibility, $\beta = -22.06$, $t(53) = 1.58$, $p = .12$, $R^2 = .07$. No additional main effects or interactions were significant, all $p > .80$.

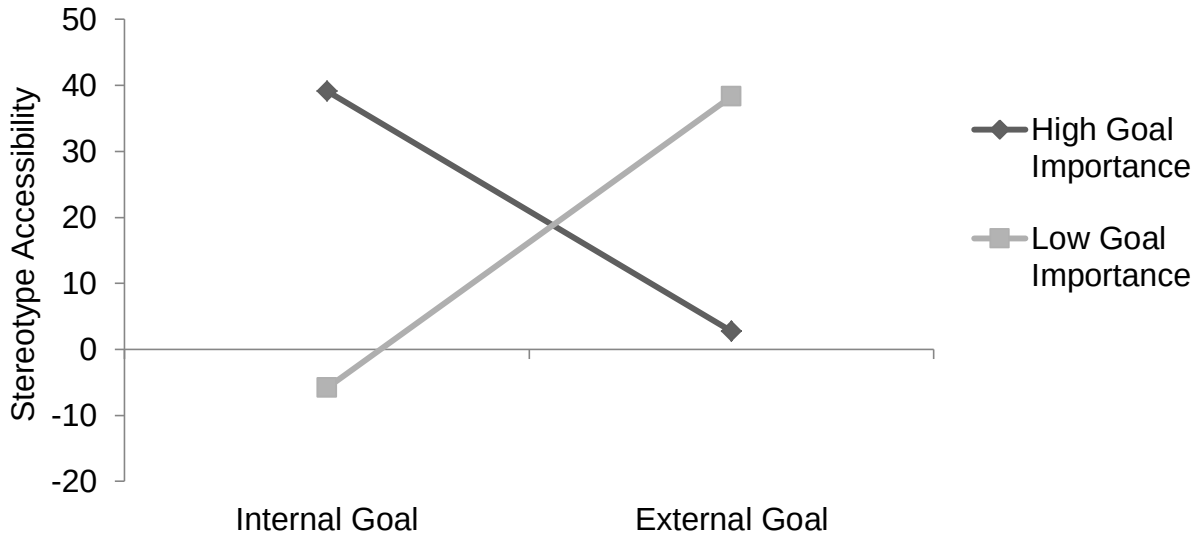

Figure 12. **Stereotype accessibility (higher scores indicate greater accessibility of Black stereotypes) as a function of goal importance and goal source, Experiment 4.**

5.2.4.3 Positive distinctiveness. To assess positive distinctiveness as a potential competing goal, the actual–ideal differentiation and the positive distinctiveness data were analysed. For the actual–ideal differentiation data, an actual–ideal differentiation score was calculated by subtracting the ideal score from the actual score, with higher scores indicating a desire for greater differentiation from the Black ethnic group (i.e., greater actual differentiation than ideal differentiation; $M = 1.98$, $SD = 1.19$). The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Actual–Ideal Differentiation standardised regression analysis. No main effects or interactions were significant, all $p > .84$; actual–ideal differentiation had no effect.

For the positive distinctiveness data, the two positive distinctiveness items were averaged into an index ($M = 4.49$, $SD = 1.40$, $\alpha = .63$). The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Positive Distinctiveness standardised regression analysis. The analysis revealed a significant main effect of positive distinctiveness, $\beta = 23.59$, $t(53) = 2.30$, $p = .026$, $R^2 = .010$; as positive distinctiveness increased, negative accessibility of Black stereotypes increased. No additional main effects or interactions were significant, all $p > .59$; goal source had no effect.

5.2.5 Additional analyses.

5.2.5.1 Goal commitment. Overall, goal commitment was moderate to high ($M = 5.53$, $SD = 1.89$). A one-way between-participants ANOVA indicated that goal commitment did not differ between the internal and external goal conditions, $F(1, 56) = 0.02$, $p = .89$, $\eta^2_p < .001$.

5.2.6 Summary. The findings of Experiment 4 demonstrated that after receiving false feedback that highlighted the potential for a goal–behaviour match (i.e., achieving the maximum score of 100% on an "egalitarianism test", indicating that there is no goal–

behaviour discrepancy), participants in both the internal and the external goal conditions exhibited accessibility of negative Black stereotypes. This pattern suggests that neither participants in the internal nor the external goal conditions controlled their stereotypes. In contrast, in Experiment 2 in which participants received false feedback that highlighted a large goal–behaviour discrepancy, participants in the internal goal condition exhibited little to no accessibility of Black stereotypes while participants in the external goal condition exhibited accessibility of Black stereotypes (see Table 3). Furthermore, the findings of Experiment 4 do not replicate the findings of Experiment 1 as expected. In Experiment 1, after reflecting on a past successful egalitarian act, as opposed to receiving false feedback that highlighted a small goal–behaviour discrepancy in Experiment 4, participants in the internal goal condition exhibited accessibility of Black stereotypes while participants in the external goal condition exhibited little to no accessibility of Black stereotypes. However, the findings of Experiment 4 do match the findings of Experiment 3. In Experiment 3, after receiving false feedback that highlighted a goal–behaviour match, participants in the internal and external goal conditions exhibited accessibility of Black stereotypes.

Table 3
Summary of stereotype accessibility findings as a function of goal source and goal–behaviour status for Experiments 1–4.

Egalitarian Goal Source	Goal–Behaviour Status	Stereotype Accessibility?	Supported Fishbach and Colleagues' Model?	Experiment #
Internal	Large (78%)	X	✓	Experiment 2
	Small (past success)	✓	✓	Experiment 1
	Small (98%)	✓	✓	Experiment 3
	Match (100%)	✓	✓	Experiment 4
External	Large (78%)	✓	✓	Experiment 2
	Small (past success)	X	✓	Experiment 1
	Small (98%)	✓	X	Experiment 3
	Match (100%)	✓	X	Experiment 4

The findings of the present experiment thus support Moskowitz and Li's (2011; see also Moskowitz, 2002) theorising that thinking about successful egalitarian acts may affirm one's identity as an egalitarian person, leading to semantic activation of the concept of "egalitarianism" and the relaxation of the egalitarian goal (i.e., indirectly suggesting that efforts to control stereotype activation cease). Indeed, in the present experiment, receiving feedback that highlighted the potential for a goal–behaviour match made participants feel completely successful. This resulted in accessibility of negative Black stereotypes, irrespective of goal source, presumably because participants "relaxed" their efforts to control stereotype activation.

In Experiments 2–4, discrepancies have not been manipulated within individual experiments. Manipulating both goal source and discrepancy size within one experiment would provide for a stronger assessment of Fishbach and colleagues' model, enabling the individual and combined influence of goal source and goal–behaviour discrepancies to be directly compared. Therefore, Experiment 5 aimed to examine how the source of an egalitarian goal (i.e., internal vs. external) and potential goal–behaviour discrepancies (i.e., goal–behaviour discrepancy vs. match) influence stereotype accessibility.

6.0 Experiment 5

Experiment 5 aimed to investigate how the source of an egalitarian goal influences stereotype accessibility when there is the potential for a goal–behaviour discrepancy or a goal–behaviour match in the context of egalitarianism. Participants were primed with either internal or external reasons to pursue an egalitarian goal. Next participants completed an "egalitarianism test" that ostensibly highlighted the potential for either a goal–behaviour discrepancy or a goal–behaviour match. Afterward, participants completed an LDT to assess stereotype accessibility.

Fishbach and colleagues' model (see Figure 1) proposes that because people with internal goals are committed to their goals, they focus on goal progress. Perceiving a large goal–behaviour discrepancy should signal that insufficient progress has been made towards the goal, resulting in increased goal adherence. In contrast, perceiving a goal–behaviour match should signal that sufficient progress has been made towards the goal, resulting in decreased goal adherence. Based on Fishbach and colleagues' reasoning, in the present experiment, participants in the internal goal condition should demonstrate stereotype accessibility when the potential for a goal–behaviour match is perceived, but little to no stereotype accessibility when the potential for a goal–behaviour discrepancy is perceived.

In contrast, Fishbach and colleagues' model (see Figure 1) proposes that because people with external goals are uncertainty committed to their goals, they focus on their goal commitment. Perceiving a large goal–behaviour discrepancy should signal that commitment is low, resulting in decreased goal adherence. In contrast, perceiving a goal–behaviour match should signal that commitment is high, resulting in increased goal adherence. Based on Fishbach and colleagues' reasoning, in the present experiment, participants in the external goal condition should demonstrate stereotype accessibility when the potential for a goal–behaviour discrepancy is perceived, but little to no stereotype accessibility when a goal–behaviour match is perceived.

Finally, the same measures of identification with the external reference group, goal importance, goal commitment, positive distinctiveness, and affect that were used in Experiments 2–4 were used in Experiment 5. The predictions for identification with the external reference group, goal importance, and goal commitment remain as per Experiments 1–4. The predictions for positive distinctiveness and affect remain as per Experiment 2 and 3.

6.1 Method

6.1.1 Participants and design. One hundred and three students[27] from the University of Birmingham completed the experiment in exchange for course credit or money (£5). The data from one participant were lost due to a computer malfunction. This left 102 participants (77 female; M_{age} = 19.7 years, *SD* = 2.21) who were randomly assigned to one of four conditions in a 2 (Goal Source: internal vs. external) × 2 (Goal–Behaviour Status: goal–behaviour discrepancy vs. goal–behaviour match) between-participants design.

6.1.2 Materials and procedure. For half of the participants, the materials and procedure in Experiment 5 were identical to the materials and procedure in Experiment 4. Importantly, these participants completed the goal–behaviour *match* framing task and received false feedback at *100%.* In contrast, for the other half of the participants, the materials and procedure in Experiment 5 were identical to the materials and procedure in Experiment 4 with one exception. Instead of completing the goal–behaviour match framing task and receiving false feedback at 100%, these individuals completed the goal–behaviour *discrepancy* framing task and received false feedback at *78%* used in Experiment 2.

6.2 Results and Discussion

6.2.1 Manipulation checks.

6.2.1.1 Instructional manipulation checks. For instructional manipulation check A, a total of 40% failed (21% failed once, 13% failed twice, 3% failed three times, 3% failed five times, and 1% failed seven times). For instructional manipulation check B, a total of 16% failed (10% failed once, 3% failed twice, 2% failed 3 times, and 1% failed five times).

[27] In Experiment 5, I collected only 25 participants per condition as I focused on the main stereotype accessibility analyses rather than the moderator analyses which have not revealed any significant findings up to this point.

6.2.1.2 Goal source primes. A goal prime index was created by averaging responses, with higher scores indicating greater agreement with the goal primes. Participants in the internal goal condition agreed moderately with the goal primes (M = 2.21, SD = 0.59, α = .65[28]); as did participants in the external goal condition (M = 2.34, SD = 0.53, α = .77).

6.2.1.3 IAT error rates. The percentage of errors that participants made during the critical blocks of the IAT was calculated (M = 3%, SD = 3.54). A two-way between-participants ANOVA indicated that the percentage of errors made during the critical blocks of the IAT did not differ depending on either the source of the goal (i.e., internal vs. external) or goal–behaviour status (i.e., goal–behaviour discrepancy vs. goal–behaviour match), all $p > .11$ (see Table 4).

Table 4
Mean Percentage and Standard Error for Errors Made During the Critical Blocks of the IAT as a Function of Internal/External Goal Source and Goal–Behaviour Discrepancy/Goal–Behaviour Match, Experiment 5.

	Internal Goal Condition	External Goal Condition
	M (SD)	
Goal–Behaviour Match	2.13 (0.70)	2.38 (0.70)
Goal–Behaviour Discrepancy	2.41 (0.68)	4.27 (0.68)

Note. Possible range = 0–64.

6.2.1.4 Goal–behaviour status.

6.2.1.4.1 Estimates prior to feedback. Participants' estimates of how well they had performed on the test and how likely they were to achieve the egalitarian goal in the future

[28] I believe that the lower internal consistency in Experiment 5 (α = .65) compared to Experiments 1 and 2 (α = .87 and α = .77, respectively) is due to some participants missing the reverse coding of one of the internal goal primes ("According to my personal values, being intolerant of Black people is ok"). Indeed, examining the *Cronbach's Alpha if Item Deleted* column indicated that removing this reversed internal goal prime would have increased the cronbach's alpha (α = .74), whereas removing any of the other internal goal primes would have decreased the cronbach's alpha. In contrast, removing this reversed internal goal prime in Experiments 1 or 2 would not have increased the cronbach's alpha.

were analysed to determine how participants interpreted their performance on the test *before* receiving the external feedback. Overall, participants indicated that they performed moderately to very well on the ostensible test (M = 5.40, SD = 1.71). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for participants' estimates of how well they performed on the test was conducted. A significant main effect of goal–behaviour status indicated that participants' perception of how well they had performed on the test was higher in the goal–behaviour match (M = 5.82, SE = 0.24) than the goal–behaviour discrepancy condition (M = 5.00, SE = 0.23), $F(1, 98) = 6.13$, $p = .015$, $\eta^2_p = .06$. No additional main effects or interactions were significant, all $p > .31$.

In addition, participants estimated that they were roughly 78% likely to achieve the egalitarian goal (M = 8.31, SD = 1.34; 8.31 was towards the higher end of choice option 71-80%). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for participants' estimates of their likely achievement of the goal was conducted. A significant main effect of goal–behaviour status indicated that participants' estimates of their likely achievement of the goal were higher in the goal–behaviour match than the goal–behaviour discrepancy condition, $F(1, 98) = 5.35$, $p = .023$, $\eta^2_p = .05$. Participants in the goal–behaviour match condition estimated their likely achievement of the goal to be roughly 81% (M = 8.62; 8.62 was just above the mid-point between choice option 71-80% and 81-90%, SE = 0.19). Participants in the goal–behaviour discrepancy condition estimated their likely achievement to be roughly 75% (M = 8.02; 8.02 was choice option 71-80%, SE = 0.18). No additional main effects or interactions were significant, all $p > .76$.

6.2.1.4.2 Interpretation of the feedback. More importantly, participants' interpretation of the feedback was examined to determine more accurately the size of the goal–behaviour discrepancy induced by the feedback. Overall, participants interpreted the feedback as indicating that they had succeeded moderately on the test (M = 2.26, SD = .78). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for participants' interpretation of the feedback was conducted. A significant main effect of goal–behaviour status indicated that participants' interpreted the feedback as demonstrating more successful performance on the test after receiving 100% (goal–behaviour match condition) compared to 78% (goal–behaviour discrepancy condition), $F(1, 98) = 97.62$, $p < .001$, $\eta^2_p = .50$. Participants who received 100% feedback (goal–behaviour match condition) indicated that they had succeeded completely on the test (M = 2.82, SE = 0.08), whereas participants who received 78% (goal–behaviour discrepancy condition) indicated that they had succeeded somewhat on the test (M = 0.73, SE = 0.08). No additional main effects or interactions were significant, all $p > .11$.

6.2.2 Affect.

6.2.2.1 Negative self-directed affect. An index of negative self-directed affect was created by averaging responses, with higher scores indicating higher negative self-directed affect (M = 0.73, SD = .1.08, α = .90). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for negative self-directed was conducted. A significant main effect of goal–behaviour status indicated that participants reported greater negative self-directed affect in the goal–behaviour discrepancy (M = 1.13, SE = 0.14) than the goal–

behaviour match condition (M = 0.32, SE = 0.14), $F(1, 98) = 16.26$, $p < .001$, $\eta^2_p = .14$. No additional main effects or interactions were significant, all $p > .47$.

6.2.2.2 Negative other-directed affect. An index of negative other-directed affect was created by averaging responses, with higher scores indicating higher negative other-directed affect (M = 0.46, SD = 1.05, $\alpha = .91$). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for negative other-directed was conducted. No main effects or interactions were significant, all $p > .15$.

6.2.2.3 Additional affect analysis. An index of general positive affect was created by averaging participants' responses to *happy* and *proud* (M = 4.47, SD = 2.08; $\alpha = .82$). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for general positive affect was conducted. A significant main effect of goal–behaviour status demonstrated that participants reported greater general positive affect in the goal–behaviour match condition (M = 5.32, SE = 0.27) than in the goal–behaviour discrepancy condition (M = 3.64, SE = 0.27), $F(1, 98) = 19.36$, $p < .001$, $\eta^2_p = .17$. No additional main effects or interactions were significant, all $p > .45$.

An index of general negative affect was created by averaging participants' responses to *threatened*, *frustrated*, *angry*, and *sad* (M = 0.43, SD = 0.74, $\alpha = .68$). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for general negative affect was conducted. A significant main effect of goal–behaviour status demonstrated that participants reported greater general negative affect in the goal–behaviour discrepancy (M = 0.66, SE = 0.10) than the goal–behaviour match condition (M = 0.20, SE = 0.10), $F(1, 98)$ =

10.75, $p = .001$, $\eta^2_p = .10$. No additional main effects or interactions were significant, all $p > .90$.

6.2.3 Stereotype accessibility. The main dependent variable was mean RTs to Black-stereotypic and stereotype-neutral words on the LDT. Three participants' data were excluded from the analyses as these participants made incorrect responses on more than 15% of trials. For the remaining 99 participants, trials with incorrect responses (3.93% of the data) and reaction times exceeding 2.5 standard deviations away from each individual participant's mean RT (2.78% of the data) were excluded from the analyses.

The data were submitted to a 2 (Goal Source: internal vs. external) × 2 (Goal–Behaviour Status: goal–behaviour discrepancy vs. goal–behaviour match) × 2 (Word Type: Black-stereotypic vs. stereotype-neutral) × 2 (Word Valence: negative vs. positive) × 2 (Block: 1 vs. 2) mixed-model ANOVA with goal source and goal–behaviour status as between-participants factors. The analysis revealed two significant main effects. First, a significant main effect of block, $F(1, 95) = 46.58$, $p < .001$, $\eta^2_p = .33$, indicated that participants responded faster in Block 2 ($M = 495$ ms, $SE = 5.46$) than Block 1 ($M = 523$ ms, $SE = 6.72$). Second, a significant main effect of word valence, $F(1, 95) = 52.28$, $p < .001$, $\eta^2_p = .36$, indicated that participants responded faster to positive words ($M = 498$ ms, $SE = 5.35$) than negative words ($M = 519$ ms, $SE = 6.49$).

The analysis also revealed a significant Word Type × Word Valence interaction, $F(1, 95) = 13.15$, $p < .001$, $\eta^2_p = .12$. Interaction means are presented in Figure 13. For positive words, participants responded faster to stereotype-neutral words than Black-stereotypic words, $t(98) = 2.06$, $p = .042$, $d = 0.13$. For negative words, however, participants responded faster to black-stereotypic words than stereotype-neutral words, $t(98) = 2.73$, $p = .007$, $d = 0.17$.

However, the theoretically important main effects of goal source, $F(1, 95) = 0.79$, $p = .38$, $\eta^2_p < .01$, and goal–behaviour status, $F(1, 95) = 0.59$, $p = .45$, $\eta^2_p < .01$, and the theoretically important interactions were all non-significant (Goal Source × Goal–Behaviour Status, $F(1, 95) = 0.06$, $p = .80$, $\eta^2_p < .01$, Goal Source × Word Type, $F(1, 95) = 0.43$, $p = .52$, $\eta^2_p < .01$, Goal–Behaviour Status × Word Type, $F(1, 95) = 1.06$, $p = .31$, $\eta^2_p = .01$, Goal Source × Goal–Behaviour Status × Word Type, $F(1, 95) = 3.15$, $p = .079$, $\eta^2_p = .03$, Goal–Behaviour Status × Word Type × Word Valence, $F(1, 95) = 1.07$, $p = .30$, $\eta^2_p = .01$, Goal Source × Word Type × Word Valence, $F(1, 95) = 0.17$, $p = .68$, $\eta^2_p < .01$, and Goal Source × Goal–Behaviour Status × Word Type × Word Valence, $F(1, 95) = 2.51$, $p = .12$, $\eta^2_p = .03$)[29]; neither goal source nor goal–behaviour status had any main or interaction effects.

[29] There was also a significant Block × Goal–Behaviour Status × Word Valence interaction, $F(1, 95) = 8.82$, $p = .004$, $\eta^2_p = .09$. As this interaction is not pertinent to the current investigation, further analysis was not conducted. No other main effects or interactions were significant, $p > .078$.

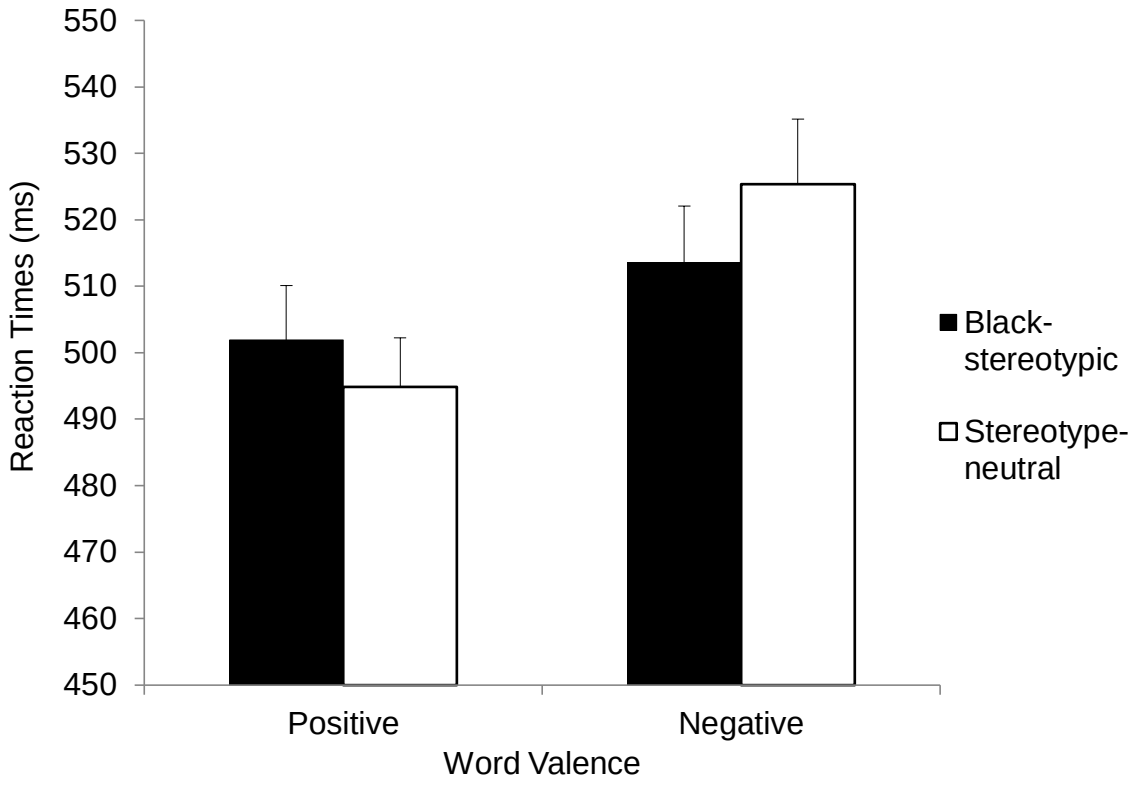

Figure 13. **Mean reaction times (ms) as a function of word type and word valence, Experiment 5.** Note. Error bars represent standard error.

6.2.4 Potential moderators. A stereotype accessibility index was created as in Experiments 2–4. Specifically, response times to Black-stereotypic words were subtracted from response times to stereotype-neutral words, with higher scores indicating stronger stereotype accessibility (*Grand M* = 9.28 ms, *SD* = 107.17).

6.2.4.1 Identification with egalitarians. An index of identification with egalitarians was created by averaging responses, with higher scores indicating higher identification (*M* = 5.20, *SD* = 1.79, α = .93[30]). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Goal–Behaviour Status (effects coded

[30] Consistent with Experiments 1–4, although one of the items in the identification measure ("I perceive myself as being similar to egalitarian people") could be considered more a measure of prototypicality than of identification, removing this item from the identification index actually reduced the reliability of the index slightly (from α = .93 to α = .92). Consequently, this item was retained in the identification index.

as -1 goal–behaviour discrepancy and +1 goal–behaviour match) × Identification standardised regression analysis. No main effects or interactions were significant, all $p > .10$; identification had no effect.

6.2.4.2 Goal importance. An index of goal importance was created by averaging responses, with higher scores indicating higher goal importance ($M = 7.02$, $SD = 1.12$, $\alpha = .88$). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Goal–Behaviour Status (effects coded as -1 goal–behaviour discrepancy and +1 goal–behaviour match) × Goal Importance standardised regression analysis. No main effects or interactions were significant, all $p > .10$; goal importance had no effect.

6.2.4.3 Positive distinctiveness. To assess positive distinctiveness as a potential competing goal, the actual–ideal differentiation and the positive distinctiveness data were analysed. For the actual–ideal differentiation data, an actual–ideal differentiation score was calculated by subtracting the ideal score from the actual score, with higher scores indicating a desire for greater differentiation from the Black ethnic group (i.e., lower actual differentiation than ideal differentiation; $M = -2.15$, $SD = 1.35$). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Goal–Behaviour Status (effects coded as -1 goal–behaviour discrepancy and +1 goal–behaviour match) × Actual–Ideal Differentiation standardised regression analysis. No main effects or interactions were significant, all $p > .07$; actual–ideal differentiation had no effect.

For the positive distinctiveness data, the two positive distinctiveness items were not indexed due to low internal consistency ($\alpha = .49$). Consequently, two standardised regression analyses (Goal Source × Goal–Behaviour Status × Item) were conducted to examine each

positive distinctiveness item separately. In neither analysis were any main effects or interactions were significant, all p > .06

6.2.5 Additional analyses.

6.2.5.1 Goal commitment. Overall, goal commitment was high (M = 6.13, SD = 1.51). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for goal commitment was conducted. No main effects or interactions were significant, all $p > .26$.

6.2.6 Summary. The findings of Experiment 5 demonstrated that all participants exhibited accessibility of negative Black stereotypes. Neither the source of the egalitarian goal nor goal–behaviour status moderated the accessibility of negative Black stereotypes. This pattern suggests that neither participants in the internal nor the external goal conditions controlled their stereotypes.

The finding that participants in both the internal and external goal conditions exhibited accessibility of negative Black stereotypes after receiving false feedback that highlighted the potential for a goal–behaviour discrepancy in Experiment 5 is inconsistent with the findings of Experiment 2 (see Table 5). Experiment 2 demonstrated that only participants in the external goal condition exhibited accessibility of Black stereotypes after perceiving the potential for a goal–behaviour discrepancy. There is one key difference between Experiment 2 and 5 which might explain the lack of replication in relation to large goal–behaviour discrepancies. In Experiment 2, participants completed a measure of their identification with egalitarians before completing the experimental manipulations (i.e., goal source and goal–behaviour status), as a means of highlighting the centrality of the egalitarian goal to participants' self-concepts. In contrast, participants' egalitarian identity was not highlighted in Experiment 5. Consequently, participants in Experiment 5 may have interpreted 78% more

positively leading them to perceive a small amount of progress toward the egalitarian goal and to exhibit accessibility of negative Black stereotypes. Indeed, although participants interpreted 78% as indicating neither success nor failure on the "test" during pilot testing, participants interpreted 78% as indicating somewhat successful performance on the "test" in Experiment 5. Therefore, while 78% certainly induced a goal–behaviour discrepancy in Experiment 5—demonstrated by participants' feeling only somewhat successful on the test, and exhibiting more negative self-directed affect after receiving feedback at 78% compared to 100%—this discrepancy may have been closer to a moderate goal–behaviour discrepancy than a large goal–behaviour discrepancy. Consequently, future research should replicate Experiment 5 with two changes: participants' egalitarian identity should be highlighted at the start of the experiment *and* the feedback should be lowered to induce a large goal–behaviour discrepancy. This will determine if a large goal–behaviour discrepancy does produce the same self-regulatory pattern as Experiment 2.

Table 5
Summary of stereotype accessibility findings as a function of goal source and goal–behaviour status for Experiments 1–5.

Egalitarian Goal Source	Goal–Behaviour Status	Stereotype Accessibility?	Supported Fishbach and Colleagues' Model?	Experiment #
Internal	Large (78%)	X	✓	Experiment 2
		✓	X	Experiment 5
	Small (past success)	✓	✓	Experiment 1
	Small (98%)	✓	✓	Experiment 3
	Match (100%)	✓	✓	Experiment 4
		✓	✓	Experiment 5
External	Large (78%)	✓	✓	Experiment 2
		✓	✓	Experiment 5
	Small (past success)	X	✓	Experiment 1
	Small (98%)	✓	X	Experiment 3
	Match (100%)	✓	X	Experiment 4
		✓	X	Experiment 5

The finding that participants in both the internal and the external goal conditions exhibited accessibility of negative Black stereotypes after receiving false feedback that highlighted the potential for a goal–behaviour match in Experiment 5 is inconsistent with the findings of Experiment 1. In Experiment 1, after reflecting on a past successful egalitarian act, as opposed to receiving false feedback that highlighted a small goal–behaviour discrepancy in Experiment 5, participants in the internal goal condition exhibited accessibility of Black stereotypes while participants in the external goal condition exhibited little to no accessibility of Black stereotypes. However, the finding that participants in both the internal and the external goal conditions exhibited accessibility of negative Black stereotypes after receiving false feedback that highlighted the potential for a goal–behaviour match is consistent with the findings of Experiments 3 and 4. Experiments 3 and 4 demonstrated that participants in both the internal and the external goal condition exhibited accessibility of negative Black stereotypes after perceiving the potential for a small goal–behaviour discrepancy (Experiment 3) or a goal–behaviour match (Experiment 4). It is also pertinent to note that participants' egalitarian identity was not highlighted at the beginning of Experiments 3 or 4 either. Fishbach and colleagues' model may only be valid in the context of egalitarianism when people's egalitarian identity is salient. Future research should replicate Experiment 4 (and potentially Experiment 3), highlighting participants' egalitarian identity at the start to determine whether the results then match Fishbach and colleagues' model and Experiment 1 of the present thesis.

The findings of Experiment 5 in relation to small goal–behaviour discrepancies thus support Moskowitz and Li's (2011; see also Moskowitz, 2002) theorising rather than Fishbach and colleagues' model. Moskowitz and Li theorised that thinking about successful egalitarian acts may affirm one's identity as an egalitarian person, leading to semantic activation of the

concept of "egalitarianism" and the relaxation of the egalitarian goal (i.e., indirectly suggesting that efforts to control stereotype activation cease). Indeed, after receiving feedback that highlighted the potential for a goal–behaviour match, participants exhibited accessibility of negative Black stereotypes, irrespective of goal source.

7.0 General Discussion

The work in the present chapter aimed to examine how the source of an egalitarian goal (internal vs. external) influences stereotype accessibility when there is a large goal–behaviour discrepancy, a small goal–behaviour discrepancy, or a goal–behaviour match in the context of egalitarianism.

7.1 Stereotype Accessibility

7.1.1 Internal goals. The findings of Experiment 2 suggested that participants in the internal goal condition were motivated to control their stereotypes about Black people after perceiving a lack of progress toward the egalitarian goal, consistent with egalitarian goal pursuit. These individuals demonstrated little to no accessibility of Black stereotypes after receiving false feedback that highlighted the potential for a goal–behaviour discrepancy (see Table 6). However, the findings of Experiment 5 were inconsistent with Experiment 2, failing to clarify the overall pattern of results. Furthermore, the findings of Experiment 1 and Experiments 3–5 suggested that participants in the internal goal condition were *not* motivated to control their stereotypes about Black people after perceiving progress toward the egalitarian goal, inconsistent with egalitarian goal pursuit. These individuals demonstrated accessibility of Black stereotypes after reflecting on a past successful egalitarian act (Experiment 1), and after receiving false feedback that highlighted the potential for a small goal–behaviour discrepancy or a goal–behaviour match (Experiment 3–5; Table 6).

Although perceived goal progress was not assessed and commitment was not assessed proximal to the stereotype-control opportunity (i.e., the LDT), the findings of the work in the present thesis suggest that Fishbach and colleagues' model may be valid, at least partially, in relation to internal egalitarian goals and goal–behaviour discrepancies. Fishbach and colleagues' model proposes that people with internally-generated goals focus on progress. Perceiving a large goal–behaviour discrepancy signals a lack of progress towards the goal, and thereby increases goal adherence so that progress can be made. In contrast, perceiving a small goal–behaviour discrepancy signals sufficient progress has been made towards the goal, which leads to the "relaxation" of goal pursuit, and thereby decreases goal adherence.

7.1.2 External goals. The findings of Experiments 2–5 suggested that participants in the external goal condition were *not* motivated to control their stereotypes about Black people after perceiving a lack of progress toward the egalitarian goal, or after progress toward/completion of the egalitarian goal, inconsistent with egalitarian goal pursuit. These individuals demonstrated accessibility of (predominantly negative) Black stereotypes after receiving false feedback that highlighted the potential for a goal–behaviour discrepancy (Experiments 2 and 5; see Table 6), and after receiving false feedback that highlighted the potential for a small goal–behaviour discrepancy or a goal–behaviour match (Experiment 3–5; see Table 6).

Although perceived goal progress was not assessed and commitment was not assessed proximal to the stereotype-control opportunity (i.e., the LDT), the findings of the work in the present thesis suggest that Fishbach and colleagues' model may *not* be valid in relation to external egalitarian goals and goal–behaviour discrepancies. Fishbach and colleagues' model proposes that people with externally-imposed goals focus on commitment. Perceiving a large goal–behaviour discrepancy signals low commitment to the goal, and thereby decreases goal

adherence. In contrast, perceiving a small goal–behaviour discrepancy signals high commitment to the goal, and thereby increases goal adherence.

The findings of Experiment 1, however, contradict both Experiments 3–5 and Fishbach and colleagues' model for external goals. Participants who were primed with external reasons to act egalitarian and reflected on a past successful egalitarian act that arose from social or political pressure in Experiment 1, exhibited little to no accessibility of negative Black stereotypes. While the findings of Experiments 2–5 suggest that discrepancy size may not matter for external goals, the finding of Experiment 1 suggests that there may be circumstances where discrepancy size does matter for external goals.

Table 6
Summary of stereotype accessibility findings as a function of goal source and goal–behaviour status, Experiments 1–5.

Egalitarian Goal Source	Goal–Behaviour Status	Stereotype Accessibility?	Supported Fishbach and Colleagues' Model?	Experiment #
Internal	Large (78%)	X	✓	Experiment 2
		✓	X	Experiment 5
	Small (98%)	✓	✓	Experiment 3
	Match (100%)	✓	✓	Experiment 4
		X	X	Experiment 5
External	Large (78%)	✓	✓	Experiment 2
		✓	✓	Experiment 5
	Small (98%)	✓	X	Experiment 3
	Match (100%)	✓	X	Experiment 4
		✓	X	Experiment 5

7.1.3 Limitations

A limitation of Experiments 1–4 is that the size of the goal-behaviour discrepancy is manipulated between experiments (i.e., a small goal–behaviour discrepancy was induced in Experiment 1 and Experiments 3 and 4, whereas a large goal–behaviour discrepancy was induced in Experiment 2). As this thesis aimed to examine whether Fishbach and colleagues'

model is valid in relation to egalitarian goals and stereotype activation, it is important to determine whether the effect of goal source is moderated by goal–behaviour discrepancy size (see Table 7 for the mean reaction times and standard errors for word type as a function of goal source and goal–behaviour discrepancy size). As a result, I analysed equivalent conditions across Experiments 2–5, grouping the large goal–behaviour discrepancy manipulation in Experiments 2 and 5 together, and the small goal–behaviour discrepancy manipulations in Experiments 3–5 together, and including Experiment as a covariate to control for multiple experiments (see Appendix G for the full analysis). The analysis revealed that participants in the internal goal condition were motivated to control their stereotypes about Black people; these individuals demonstrated little to no accessibility of Black stereotypes. Participants in the external goal condition, however, were not motivated to control their stereotypes about Black people; these individuals demonstrated accessibility of Black stereotypes. However, while the size of the goal-behaviour discrepancy did not interact with either the source of the goal or the type of word, caution must be taken when interpreting the results of this analysis because goal–behaviour status was only manipulated within Experiment 5. Thus, even though these results may appear to suggest that goal source is a more important determinant of self-regulation than is self-regulatory success versus failure, these results cannot be taken as adjudicating unequivocally on the relative impact of the two factors.

Table 7

Mean Reaction Times (ms) and Standard Error as a function of Internal/External Goal Source, Goal–Behaviour Status, Experiment Number, and Word Type.

Egalitarian Goal Source	Goal–Behaviour Status	Experiment #	Word Type			
			Negative Black-Stereotypic	Positive Black-Stereotypic	Negative Stereotype-Neutral	Positive Stereotype-Neutral
			M (*SD*)			
Internal	Large (78%)	Experiment 2	508 (9.20)	496 (8.41)	526 (11.50)	489 (7.70)
		Experiment 5	512 (13.22)	502 (12.95)	529 (13.89)	489 (10.23)
	Small (98%)	Experiment 3	534 (10.91)	526 (11.34)	554 (11.33)	517 (9.99)
	Match (100%)	Experiment 4	522 (12.41)	509 (11.67)	531 (12.12)	497 (10.65)
		Experiment 5	532 (16.15)	509 (11.00)	531 (12.21)	507 (12.27)
External	Large (78%)	Experiment 2	501 (6.85)	486 (6.50)	525 (7.99)	495 (7.92)
		Experiment 5	510 (9.39)	500 (8.76)	521 (10.46)	487 (7.13)
	Small (98%)	Experiment 3	540 (12.29)	524 (15.01)	561 (14.84)	519 (12.07)
	Match (100%)	Experiment 4	501 (10.29)	495 (11.37)	513 (14.12)	485 (9.83)
		Experiment 5	503 (10.50)	498 (11.81)	526 (11.51)	499 (8.73)

7.2 Identification with Egalitarians

Past research has demonstrated that identification with the group imposing the external goal is a key moderator of goal adherence (see Fishbach et al., 2011; Sassenberg, 2011). For example, Sassenberg et al. (2011) examined the influence of goal–behaviour discrepancies in relation to ingroup norms that are considered part of the self-concept on affect and motivation. They demonstrated that receiving false feedback highlighting a mismatch between the participant's performance and the ingroup norm (being a good co-student (Experiment 2) or cognitive flexibility (Experiment 3)) resulted in greater negative affect (Experiment 2) and greater motivation to reduce the discrepancy (Experiment 3) than receiving false feedback highlighting a match between participant's performance and the ingroup norm. Sassenberg et al. concluded that ingroup norms are adopted as peoples own personal standards, hence the self-regulatory pattern for ingroup norms and internal goals are identical.

However, in the present experiment, the influence of goal source on stereotype accessibility was not moderated by identification with the external reference group (in this case, egalitarians). As "egalitarianism" may not be a concept that participants reflect on regularly, participants may not have previously thought of themselves as a member of the egalitarian group. Consequently, the egalitarian group may lack entitativity, "the degree of having the nature of an entity, of having a real existence" (Campbell, 1958, p.17). Campbell proposed three key factors that influence the entitativity of a group: common fate (the extent to which group members experience similar outcomes), similarity (the extent to which group members are similar to one another in terms of their appearance and their actions), and proximity (the extent to which group members are spatially close to one another). Entitativity is important for identification as people identify more highly with highly entitative groups

compared to lowly entitative groups (Hogg, Sherman, Dierselhuis, Maitner, & Moffitt, 2007; Sacchi, Castano, & Brauer, 2009).

7.3 Conclusion

In conclusion, Chapter 3 aimed to investigate the effect of internal and external egalitarian goals on stereotype accessibility following false feedback that manipulated the size of the goal–behaviour discrepancy. The work in this chapter demonstrated that for participants in the internal goal condition, highlighting the potential for a goal–behaviour discrepancy resulted in little to no accessibility of Black stereotypes (Experiment 2). In contrast, highlighting the potential for a small goal–behaviour discrepancy (Experiment 3) or a goal–behaviour match (Experiments 4 & 5) resulted in accessibility of negative Black stereotypes. For participants in the external goal condition, highlighting the potential for a goal–behaviour discrepancy (Experiment 2), a small goal–behaviour discrepancy (Experiment 3), or a goal–behaviour match (Experiments 4 & 5) all resulted in accessibility of negative Black stereotypes. These findings have important implications for our understanding of the impact of goal–behaviour discrepancies on self-regulation for internal and external egalitarian goals.

CHAPTER 4

CONSEQUENCES OF COMPLIANCE VERSUS NONCOMPLIANCE WITH INTERNAL OR EXTERNAL EGALITARIAN GOALS FOR SOCIAL CATEGORISATION

In this chapter, I turned to social categorisation to determine how early egalitarian goals take effect. In Experiment 6, I examined the effect of goal–behaviour discrepancies on social categorisation for individuals with either an internal or external goal to act egalitarian. Specifically, participants were primed with either internal reasons (personal beliefs and values) or external reasons (other people's expectations) to act in an egalitarian manner before receiving false feedback from an "egalitarianism test" that ostensibly highlighted the potential for a goal–behaviour discrepancy or a goal–behaviour match (i.e., no goal–behaviour discrepancy). Participants completed a categorisation task assessing how they categorised racially ambiguous faces and how quickly they categorised racially ambiguous and unambiguous faces. Participants in the internal goal condition who received false feedback that highlighted the potential for a goal–behaviour discrepancy categorised racially ambiguous faces as quickly as unambiguous White faces, and categorised unambiguous Black faces faster than both racially ambiguous faces and unambiguous White faces. This finding suggests that these individuals were motivated to demonstrate their inclusiveness. In contrast, participants in the internal goal condition who received false feedback that highlighted the potential for a goal–behaviour match, and participants in the external goal condition irrespective of the discrepancy manipulation, were slower to categorise racially ambiguous faces than unambiguous faces (whether unambiguous Black or White faces). This finding suggests that these individuals were motivated to be accurate in their categorisations of racially ambiguous faces.

1.0 General Introduction

Influential social-cognitive psychologists (e.g., Allport, 1954; Brewer, 1988; Fiske & Neuberg, 1990) have argued that social categorisation is a key aspect of person perception. Allport (1954), for example, argued that social categorisation is an automatic and inevitable consequence of person perception and results in the automatic activation of relevant social information like stereotypes. As a result of this theoretical reasoning, both Brewer's (1998) dual-process model and Fiske and Neuberg's (1990; see also Fiske, Lin, & Neuberg, 1999) continuum model propose that when social perceivers are exposed to a person, they automatically and inevitably categorise the person based on perceptual properties such as gender, age, and ethnicity (Allport, 1954; Brewer, 1988; Fiske, 1998; Fiske & Neuberg, 1990).

In more recent years, researchers have argued that social categorisation is more malleable than previously thought (Bodenhausen & Macrae, 1998; Livingston & Brewer, 2002; Macrae, Bodenhausen, & Milne, 1995; Macrae, Quinn, Mason, & Quadflieg, 2005; Pendry & Macrae, 1996; Quinn & Macrae, 2005). Research has demonstrated that social categorisation is moderated by people's current processing goals (e.g., Livingston & Brewer, 2002; Macrae et al. 1995; Macrae et al., 2005; Pendry & Macrae, 1996; Quinn & Macrae, 2005). Quinn and Macrae (2005; see also Macrae et al., 2005), for example, used a repetition priming task to assess automatic social category activation. During the initial encoding phase, half the participants categorised faces according to sex (active-encoding condition) and half viewed the faces (passive-encoding condition). During the test phase, all participants categorised the faces presented during the encoding phase, plus an equal number of new faces, according to sex. If social categorisation is automatic, then all participants should respond faster to repeated faces (vs. new faces) regardless of encoding condition. This effect

would be due to a residual memory trace being established when faces are processed during the encoding phase, which facilitates responding when a face is reencountered. Contrary to this, Quinn and Macrae found no such facilitation to repeated faces (vs. new faces) in the passive-encoding condition. In contrast, participants were faster to categorise repeated faces compared to new faces in the active-encoding condition. Therefore, in the absence of a goal to engage in social categorisation, social categorisation did not occur, suggesting social categorisation is not automatic but, instead, is contingent on processing goals.

Indeed, Livingston and Brewer (2002) argued that when a social perceiver has a categorisation goal, they categorise targets according to the social groups to which they belong (category-based processing). In the absence of a categorisation goal, however, a social perceiver processes only the perceptual features of the targets (cue-based processing). Livingston and Brewer (Experiments 1 & 4) found that participants who simply viewed White faces and Black faces that were high in prototypicality (i.e., more typical Black facial features) and low in prototypicality (i.e., fewer typical Black facial features) exhibited cue-based processing; that is, they evaluated the faces based on their affective reactions to the perceptual features of the faces. Specifically, these participants exhibited greater negativity to Black faces high in prototypicality compared to Black faces low in prototypicality. In contrast, Livingston and Brewer (Experiment 4) found that participants who mentally categorised White faces and Black faces that were high and low in prototypicality according to their ethnicity exhibited category-based processing; that is, they evaluated the faces based on social information about the category to which the faces belonged. Specifically, the prototypicality effect (i.e., greater negativity to Black faces high in prototypicality compared to Black faces low in prototypicality) disappeared, and, instead, Black faces high and low in

prototypicality were evaluated equally negatively. This research indicates that social categorisation is contingent upon a categorisation goal.

There has been less consideration of whether higher-order goals, like the goal to act egalitarian, influence social categorisation. Yet, in order to determine *how* goals influence stereotype activation and control, it is important to establish *when* goals take effect. It may be that egalitarian goals do not influence social categorisation at all. Egalitarian goals may only be activated when there is some implicit or explicit recognition that stereotyping might occur (e.g., Wilson & Brekke, 1994). For example, when a person is encountered, we may initially process the perceptual properties of the face (e.g., dark skin; cue-based processing), which signals the category membership of the person (e.g., Black; category-based processing) and then activates associated stereotypes (e.g., aggressive; Brewer, 1988; Fiske & Neuberg, 1990; Livingston & Brewer, 2002). The act of social categorisation may lead the social perceiver to recognise the potential for stereotyping the target, resulting in the activation of the egalitarian goal which inhibits the activated Black-stereotype. Moskowitz and colleagues' research (Moskowitz et al., 1999; Moskowitz et al., 2000; see Moskowitz & Ignarri, 2009, for a review) has demonstrated that once a category is activated, stereotype control or inhibition is exhibited by people with chronic egalitarian goals and temporary internal egalitarian goals (Moskowitz et al., 1999; Moskowitz et al., 2000; Moskowitz & Li, 2011; Moskowitz & Stone, 2012; see Moskowitz & Ignarri, 2009, for a review).

Alternatively, egalitarian goals may affect whether targets at this stage are categorised as members of the stereotyped group. When a social perceiver processes the perceptual properties of a target's face, activation of an egalitarian goal may occur for those social perceivers who repeatedly and consistently act egalitarian in the presence of such perceptual properties (Bargh, 1990; Bargh & Chartrand, 1999). For example, dark skin may trigger the

egalitarian goal if the social perceiver has repeatedly and consistently acted in an egalitarian manner towards Black people. In this case, awareness of possible stereotyping would not be needed to initiate the egalitarian goal. This is likely to occur for people who have a chronic or internal egalitarian goal, as these individuals pursue the egalitarian goal frequently and have established an implicit association between a group and the egalitarian goal, but not for people who have a nonchronic or external egalitarian goal, as these individuals value egalitarianism but have yet to establish an implicit association between a group and the egalitarian goal (Hausmann & Ryan, 2004; Moskowitz et al., 1999; Moskowitz et al., 2000; see Moskowitz & Ignarri, 2009, for a review).

Blascovich, Wyer, Swart, and Kibler (1997) were one of the first to examine *how* a prejudiced or unprejudiced attitude, which might reflect underlying chronic motivation, influences social categorisation of ambiguous faces[31]. They argued that prejudiced individuals, who are presumably strongly identified with their ingroup (i.e., they view the ingroup more positively than the outgroup), are motivated to accurately categorise ambiguous faces in order to protect the ingroup status. In contrast, nonprejudiced individuals, who are presumably lower in identification with their ingroup (i.e., they view both the ingroup and outgroup positively), are not motivated to accurately categorise ambiguous faces because accidentally including an outgroup member in the ingroup is not perceived as a threat to the ingroup status. Indeed, Blascovich et al. found that the slowdown when categorising ambiguous faces compared to unambiguous faces was significantly larger for prejudiced participants than for unprejudiced participants. This supports the notion that prejudiced

[31] Examining how people categorise racially ambiguous faces is becoming ever more important as the number of multi-racial people within the UK continues to grow (1.4% in 2001 UK Census versus 2.2% in the 2011 UK Census (Office of National Statistics, 2011). Whether a racially ambiguous face is categorised as an ingroup member or as an outgroup member has important implications for whether subsequent stereotypes are activated and applied.

individuals are more motivated than nonprejudiced individuals to accurately categorise ambiguous faces (or at least "accurately" in relation to intergroup differentiation goals).

Extending Blascovich and colleagues' (1997) investigation, Castano, Yzerbyt, Bourguignon, and Seron (2002) examined how participants categorise ambiguous faces as well as how long they spend categorising ambiguous faces. The analysis revealed that participants low in identification with the ingroup spent longer categorising ambiguous faces in comparison to less ambiguous faces, regardless of whether the less ambiguous faces were ingroup members (Northern Italian) or outgroup members (Southern Italian). Additionally, these participants were equally likely to categorise ambiguous faces as "Northern Italian" or "Southern Italian". Castano et al. concluded that participants who were low in identification with the ingroup were motivated to accurately categorise ambiguous targets. In contrast, participants high in identification with the ingroup spent longer categorising faces as the likelihood of ingroup membership increased; that is, as the percentage of ingroup member features in the morphed faces increased, categorisation times also increased. Additionally, these participants were more likely to categorise ambiguous faces as "Southern Italian" outgroup members than "Northern Italian" ingroup members, indicative of the ingroup overexclusion effect (cf. Leyens & Yzerbyt, 1992). Castano et al. concluded that this overall pattern was more consistent with a motivation to defend the integrity of the ingroup rather than purely an accuracy motivation.

Recent research by Chen, Moons, Gaither, Hamilton, and Sherman (2014, Experiment 4) has demonstrated that social categorisation is also influenced by motivation. In their research, participants categorised unambiguous White and Black faces, and racially ambiguous faces that varied in the degree of ambiguity, according to race. Importantly, participants were given the opportunity to categorise the faces as "White", "Black", or

"multiracial". The results demonstrated that higher internal motivation to act nonprejudiced was associated with a higher likelihood of categorising racially ambiguous faces as multiracial, and that this relationship increased in strength as the degree of racial ambiguity in the faces increased. There was also a tendency for higher external motivation to act nonprejudiced to be associated with a lower likelihood of categorising racially ambiguous faces as multiracial, but this was not related to the degree of racial ambiguity. Chen and colleagues concluded that people high in internal motivation engage in individuated processing in order to accurately categorise racially ambiguous faces. They also tentatively concluded that people high in external motivation use the multiracial category sparingly even when racial ambiguity is high, perhaps because they fear the term is not politically correct and therefore socially unacceptable.

The experiment in the present chapter examined how the source of an egalitarian goal (i.e., internal vs. external) affects social categorisation of racially ambiguous faces into the ingroup or the outgroup. In line with the previous research in this thesis, I also examined whether goal–behaviour discrepancies (i.e., goal–behaviour discrepancy or a goal–behaviour match) effect social categorisation. Although research has examined how prejudice level influences social categorisation of racially ambiguous faces into the ingroup or outgroup (e.g., Blascovich et al., 1997; Castano et al., 2002), motivation may influence social categorisation differently than prejudice level. Plant and Devine (1998) investigated the convergent validity of their internal and external motivation to respond without prejudice scales with measures of prejudice. Although they found a strong negative correlation between internal motivation and measures of prejudice, there was only a small negative correlation between external motivation and measures of prejudice. To the best of my knowledge, research has not yet

examined the *how* internal or external motivation influences social categorisation of racially ambiguous faces into the ingroup or the outgroup.

2.0 Experiment 6

Experiment 6 examined how the source of an egalitarian goal (i.e., internal vs. external) influences social categorisation when there is the potential for a goal–behaviour discrepancy or a goal–behaviour match in the context of egalitarianism. Participants were either primed with internal reasons or with external reasons to pursue an egalitarian goal. Next participants completed an "egalitarianism test" that ostensibly highlighted either the potential for a goal–behaviour discrepancy or a goal–behaviour match. Afterward, participants completed a speeded categorisation task where they indicated the race of unambiguous White targets, unambiguous Black targets, and racially ambiguous targets.

All of the predictions throughout this thesis have been based on Fishbach and colleagues' model. However, it is not clear whether Fishbach and colleagues' model is valid in relation to social categorisation. Their model predicts the conditions that should result in increased or decreased goal adherence (e.g., an internal goal and a large goal–behaviour discrepancy result in increased goal adherence). Although it seems straightforward to argue that adhering to an egalitarian goal would involve stereotype control/inhibition, as I have suggested throughout this thesis, it is less obvious whether individuals pursuing an egalitarian goal would view avoiding categorisation as part of this goal. It is possible that being egalitarian might be interpreted as being colour-blind and ignoring racial category membership altogether, but it should be equally possible that being egalitarian might be interpreted as being multicultural and paying attention to racial category membership (Moskowitz & Ignarri, 2009; Richeson & Nussbaum, 2004). Thus, the applicability of Fishbach and colleagues' model to social categorisation remains unclear. Moreover, given the

unclear relation between goal source (internal/external) and prejudice level (low/high), the extant literature on how prejudice level affects categorisation offered only tentative insights. As a result, Experiment 6 was primarily exploratory and no specific predictions were made.

2.1 Method

2.1.1 Participants and design. Seventy-nine White students[32] from the University of Birmingham completed the experiment in exchange for course credit or money (£4). Fourteen participants' data were removed from data analysis for having participated in one of the previous experiments reported in this thesis. This left 65 participants (55 female; M_{age} = 19.1 years, *SD* = 0.90) who were randomly assigned to one of four conditions in a 2 (Goal Source: internal vs. external) × 2 (Goal–Behaviour Status: goal–behaviour discrepancy vs. goal–behaviour match) between-participants design.

2.1.2 Materials and procedure. The majority of materials and procedures were identical to those used in previous experiments in the present thesis except for the categorisation measure.

2.1.2.1 Instructional manipulation check A. The same instructional manipulation check A from Experiments 1–5 was used in Experiment 6.

2.1.2.2 Goal source manipulation. Next, participants completed the goal source manipulation from Experiments 2–5. Specifically, participants either rated their agreement with 10 internal reasons to act egalitarian towards Black people (internal goal condition) or 10 external reasons to act egalitarian towards Black people (external goal condition).

[32] Based on past research that has collected less than 20 participants per condition when examining the effect of prejudice-level (Blascovitch et al., 1997) or identification-level (Castano et al., 1997), I determined that collecting 20 participants per condition to examine the effect of goal source and goal–behaviour status would be sufficient in Experiment 6. I did not collect more than 20 participants per condition because I focused on the main categorisation analyses rather than the moderator analyses which had not revealed any significant findings up to this point.

2.1.2.3 Goal–behaviour discrepancy framing task. Participants either completed the large goal–behaviour discrepancy framing task used in Experiments 2 and 5 (78%) or the goal–behaviour match framing task used in Experiments 4 and 5 (100%). In the goal–behaviour discrepancy framing task, participants learned that they would be completing a test of how motivated they are to act egalitarian and would receive feedback about their performance, based on how *quickly and accurately* they responded. The test was actually an IAT designed to heighten participants' awareness of their own potential for nonegalitarian responding. Specifically, participants only completed prejudice-*congruent* trials to ensure prejudicial responding was salient; the need for participants to respond quickly and accurately on the test was emphasised in order to inflate error rates; finally, error feedback was provided to increase participants' awareness of their errors during the test.

In the goal–behaviour match framing task, participants learned that they would be completing a test of how motivated they are to act egalitarian and would receive feedback about their performance, based on how *accurately but not how quickly* they responded. The test was actually an IAT designed to heighten participants' awareness of their own potential for egalitarian responding. Specifically, participants only completed prejudice-*incongruent* trials to ensure egalitarian responding was salient; the need for participants to respond accurately but not quickly on the test was emphasised in order to reduce error rates; finally, correct-response feedback was provided to increase participants' awareness of their correct responses during the test.

2.1.2.3 Experience during the test. After completing the discrepancy framing task, participants waited 20 seconds while the computer ostensibly calculated their test score. Participants completed the same estimations about their performance from Experiments 2–5.

2.1.2.4 Feedback. Participants were presented with a graph depicting their performance. In the goal–behaviour discrepancy condition, the graph depicted an arrow extending upwards to 78% on a scale from 0–100%, as in Experiments 2 and 5. In the goal–behaviour match condition, the graph depicted an arrow extending upwards to 100% on a scale from 0–100%, as in Experiments 4 and 5. The accompanying wording was consistent with each goal source.

2.1.2.5 Interpretation of feedback. As per Experiments 4 and 5, immediately after receiving the false feedback, participants indicated how well they had performed on the test, on a 7-point scale ranging from -3 (*completely failed*) to +3 (*completely succeeded*). Additionally, participants indicated how well the feedback they received matched their expectations, on a 7-point scale from 0 (*not at all)* to 8 (*completely)*.

2.1.2.6 Affect measures. The same affect measures from Experiments 2–5 were used in the present experiment.

2.1.2.7 Categorisation task. Participants learned that the experiment would be assessing the ease with which they extract information about ethnicity from faces. Participants were presented with 39 randomly ordered faces[33] (13 unambiguous White faces, 13 unambiguous Black faces, and 13 racially ambiguous faces) superimposed on a black

[33] The ambiguous faces were created using MorphMan (STOIK Imaging, 2000). One unambiguous Black face and one unambiguous White face were morphed at 50% so that the racially ambiguous face contained 50% of the unambiguous White face and 50% of the unambiguous Black face. Afterward, the faces (racially ambiguous and unambiguous faces) were pilot tested to ensure that the racially ambiguous faces were truly racially ambiguous and the unambiguous faces were truly unambiguous. Out of the 35 racially ambiguous faces originally created, 13 were selected because participants were equally likely to choose "Black" or "White" when categorising these faces (based on a binomial test, these 13 faces did not significantly differ from 50%; all $p > .36$). For the unambiguous faces (i.e., the 13 unambiguous Black faces & 13 unambiguous White faces used to create the 13 useable ambiguous faces), "Black" responses were coded as 1 and "White" responses were coded as 0. Unambiguous Black faces were categorised as Black ($M = .97$, $SD = 0.16$) and unambiguous White faces were categorised as White ($M = .01$, $SD = 0.11$), indicating that the unambiguous faces were prototypical of each category.

background (see Figure 14). For each face, participants indicated whether the target's ethnicity was Black or White, using the "/" and the "Z" keys; the correspondence between the keys and the ethnicities was randomised across participants. On each trial, participants were presented with a fixation cross for a random amount of time ranging from 250 ms to 750 ms, followed by a face that remained on screen until participants made a response[34].

Figure 14. **Sample facial stimuli used in the categorisation task, Experiment 6.** An example of an unambiguous Black face (far left), an unambiguous White face (far right) and a racially ambiguous face (middle) with 50% of each unambiguous face.

2.1.2.8 **Goal commitment.** Participants completed the measure of goal commitment from Experiments 2–5.

Lastly, participants provided demographic information and underwent "process" debriefing (McFarland et al., 2007) to mitigate any lasting effects of the negative feedback.

[34] Participants also completed a series of exploratory measures in Experiment 6. The measure of goal importance used in Experiments 2–5 and the measure of self-concordance used in Experiment 4 were also presented along with the goal commitment measure in a random order. Additionally, following the measure of goal commitment, participants completed the measure of identification with egalitarians from Experiments 2–5, and then the measure of contact with Black people and reactance from Experiments 2–5 in a random order. As the results of the categorisation task were only marginally significant, I did not analyse the exploratory measures.

2.2 Results

2.2.1 Manipulation checks.

2.2.1.1 Instructional manipulation check A. For instructional manipulation check A, a total of 31% failed (9% failed once, 9% failed twice, 8% failed three times, 2% failed four times, and 3% failed five times).

2.2.1.2 Goal source primes. A goal prime index was created by averaging responses, with higher scores indicating greater agreement with the goal primes. Participants in the internal goal condition agreed moderately with the goal primes (M = 2.38, SD = 0.51, α = .58[35]); as did participants in the external goal condition (M = 2.24, SD = 0.62, α = .81).

2.2.1.3 IAT error rates. The percentage of errors that participants made during the critical blocks of the IAT was calculated (M = 2%, SD = 2.91). A two-way between-participants ANOVA indicated that the percentage of errors made during the IAT did not differ depending on either the source of the goal (i.e., internal vs. external) or goal–behaviour status (i.e., goal–behaviour discrepancy vs. goal–behaviour match), all $p > .12$ (see Table 8).

Table 8
Mean Percentage and Standard Error for Errors Made During the Critical Blocks of the IAT as a Function of Internal/External Goal Source and Goal–Behaviour Discrepancy/Goal–Behaviour Match, Experiment 6.

	Internal Goal Condition	External Goal Condition
	M (*SD*)	
Goal–Behaviour Match	1.25 (0.75)	2.71 (0.67)
Goal–Behaviour Discrepancy	2.34 (0.68)	3.13 (0.81)

Note. Possible range = 0–64.

[35] I believe that the lower internal consistency in Experiment 6 (α = .63) compared to Experiments 1 and 2 (α = .87 and α = .77, respectively) is due to some participants missing the reverse coding of one of the internal goal primes ("According to my personal values, being intolerant of Black people is ok"). Indeed, examining the *Cronbach's Alpha if Item Deleted* column indicated that removing this reversed internal goal prime would have increased the cronbach's alpha (α = .70), whereas removing any of the other internal goal primes would have had no effect or decreased the cronbach's alpha.

2.2.1.4 Goal–behaviour status.

2.2.1.4.1 Estimates prior to feedback. Participants' estimates of how well they had performed on the test and how likely they were to achieve the egalitarian goal in the future were analysed to determine how participants interpreted their performance on the test *before* receiving the external feedback. Overall, participants indicated that they performed moderately to very well on the ostensible test (M = 5.12, SD = 1.88). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for participants' estimates of how well they performed on the test was conducted. A significant main effect of goal–behaviour status indicated that participants' perception of how well they had performed on the test was higher in the goal–behaviour match (M = 5.53, SE = 0.32) than the goal–behaviour discrepancy condition (M = 4.58, SE = 0.33), $F(1, 61) = 4.24$, $p = .044$, $\eta^2_p = .07$. No additional main effects or interactions were significant, all $p > .20$.

In addition, participants estimated that they were roughly 77% likely to achieve the egalitarian goal (M = 8.18, SD = 1.40). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for participants' estimates of their likely achievement of the goal was conducted. No main effects or interactions were significant, all $p > .075$.

2.2.1.4.2 Interpretation of the feedback. More importantly, participants' interpretation of the feedback was examined to determine more accurately the size of the goal–behaviour discrepancy induced by the feedback. Overall, participants interpreted the feedback as indicating that they had completely succeeded on the test (M = 3.00, SD = 1.13). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for participants'

interpretation of the feedback was conducted. A significant main effect of goal–behaviour status indicated that participants' interpreted the feedback as demonstrating more successful performance on the test after receiving 100% (goal–behaviour match condition) compared to 78% (goal–behaviour discrepancy condition), $F(1, 61) = 22.94$, $p < .001$, $\eta^2_p = .27$. Participants who received 100% feedback indicated that they had succeeded completely on the test ($M = 2.58$, $SE = 0.17$), whereas participants who received 78% indicated that they had succeeded somewhat on the test ($M = 1.40$, $SE = 0.18$). No additional main effects or interactions were significant, all $p > .37$.

In addition, participants indicated that their estimates matched their feedback moderately to very well ($M = 5.29$, $SD = 2.01$). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for participants' interpretation of the feedback was conducted. A significant main effect of goal–behaviour status indicated that the feedback matched participants' expectations in the goal–behaviour match condition ($M = 5.90$, $SE = 0.34$) to a greater extent than in the goal–behaviour discrepancy condition ($M = 4.63$, $SE = 0.36$), $F(1, 61) = 6.62$, $p = .013$, $\eta^2_p = .10$. No additional main effects or interactions were significant, all $p > .68$.

2.2.2 Affect.

2.2.2.1 Negative self-directed affect. An index of negative self-directed affect was created by averaging responses, with higher scores indicating higher negative self-directed affect ($M = 0.76$, $SD = .1.02$, $\alpha = .90$). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for negative self-directed was conducted. The analysis revealed a significant main effect of goal–behaviour status, indicating that participants

reported greater negative self-directed affect in the goal–behaviour discrepancy (M = 1.08, SE = 0.18) than the goal–behaviour match condition (M = 0.54, SE = 0.17), $F(1, 61) = 4.80$, p = .032, $\eta^2_p = .07$.

Additionally, the analysis revealed a significant Goal Source × Goal–Behaviour Status interaction, $F(1, 61) = 4.67$, $p = .035$, $\eta^2_p = .07$. In the internal goal condition, participants reported equally low levels of negative self-directed affect in the goal–behaviour discrepancy (M = 0.74, SE = 0.93) and the goal–behaviour match conditions (M = 0.73, SE = 1.02), $t(31) = 0.02$, p = .98, $d < 0.01$. In the external goal condition, however, participants reported greater negative self-directed affect in the goal–behaviour discrepancy (M = 1.41, SE = 0.41) than the goal–behaviour match condition (M = 0.34, SE = 0.10), $t(13.32) = 2.54$, p = .024, $d = 0.98$[36].

2.2.2.2 Negative other-directed affect. An index of negative other-directed affect was created by averaging responses, with higher scores indicating higher negative other-directed affect (M = 0.46, SD = 0.69, α = .81). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for negative other-directed was conducted. No main effects or interactions were significant, all $p > .28$.

2.2.2.3 Additional affect analysis. An index of general positive affect was created by averaging participants' responses to *happy* and *proud* (M = 4.26, SD = 1.94; α = .79). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for general positive affect was conducted. A significant main effect of goal–behaviour status demonstrated that participants reported greater general positive affect in the goal–behaviour match (M = 4.92, SE = 0.32) than the goal–behaviour discrepancy condition (M = 3.58, SE =

[36] Levene's Test for Equality of Variances was significant (F = 4.56, p = .041); therefore, a t-test not assuming equal variances was conducted.

0.34), $F(1, 61) = 8.46$, $p = .005$, $\eta^2_p = .12$. No additional main effects or interactions were significant, all $p > .20$.

An index of general negative affect was created by averaging participants' responses to *threatened*, *frustrated*, *angry*, and *sad* ($M = 0.57$, $SD = 0.70$, $\alpha = .58$). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for general negative affect was conducted. The analysis revealed a significant Goal Source × Goal–Behaviour Status interaction, $F(1, 61) = 4.38$, $p = .041$, $\eta^2_p = .07$. In the internal goal condition, participants reported similar levels of general negative affect in the goal–behaviour discrepancy ($M = 0.49$, $SE = 0.63$) and goal–behaviour match conditions ($M = 0.52$, $SE = 0.64$), $t(31) = 0.14$, $p = .89$, $d = 0.05$. However, in the external goal condition, participants reported significant higher general negative affect in the goal–behaviour discrepancy ($M = 1.04$, $SE = 0.95$) than the goal–behaviour match condition, ($M = 0.37$, $SE = 0.46$), $t(30) = 2.66$, $p = .012$, $d = 0.90$. No main effects were significant, both $p > .06$.

2.2.3 Reaction times. The first main dependent variable was mean RTs to unambiguous White faces, unambiguous Black faces, and ambiguous faces. Reaction times exceeding 2.5 standard deviations away from each individual participant's mean RT (3.44% of the data) were excluded from the analyses.

The reaction time data were submitted to a 2 (Goal Source: internal vs. external) × 2 (Goal–Behaviour Status: goal–behaviour discrepancy vs. goal–behaviour match) × 3 (Target Race: unambiguous Black vs. racially ambiguous vs. unambiguous White) mixed-model ANOVA with goal source and goal–behaviour status as between-participants factors. The analysis revealed a significant main effect of target race, $F(1.33, 60) = 82.89$, $p < .001$, $\eta^2_p = .58$[17]. Post-hoc tests using a Bonferroni correction demonstrated that participants categorised

racially ambiguous faces (*M* = 818 ms, *SE* = 35.46) significantly slower than unambiguous Black faces (*M* = 553 ms, *SE* = 14.39; *p* < .001) and unambiguous White faces (*M* = 605 ms, *SE* = 20.12; *p* < .001); they were also faster to categorise unambiguous Black faces than unambiguous White faces (*p* < .001).

The analysis also revealed two significant interactions. First, there was a significant Goal–Behaviour Status × Target Race interaction, $F(1.33, 60) = 6.22$, $p = .009$, $\eta^2_p = .09$[37]. Independent sample *t*-tests indicated that participants responded equally fast to unambiguous Black faces in the goal–behaviour discrepancy (*M* = 543 ms, *SE* = 17.85) and goal–behaviour match conditions (*M* = 564 ms, *SE* = 21.50), $t(63) = 0.75$, $p = .46$, $d = .19$. Participants also responded equally fast to unambiguous White faces in the goal–behaviour discrepancy (*M* = 603 ms, *SE* = 32.16) and goal–behaviour match conditions (*M* = 609 ms, *SE* = 24.21), $t(63) = 0.16$, $p = .88$, $d = .04$. More importantly, participants responded marginally faster to racially ambiguous faces in the goal–behaviour discrepancy (*M* = 745 ms, *SE* = 41.87) than the goal–behaviour match condition (*M* = 883 ms, *SE* = 56.57), $t(63) = 1.93$, $p = .058$, $d = .48$.

Second, there was a marginally significant Goal Source × Goal–Behaviour Status × Target Race interaction, $F(1.33, 60) = 2.83$, $p = .085$, $\eta^2_p = .04$[17]. Interaction means are presented in Figure 15. The interaction was decomposed by analysing the internal and external goal conditions separately.

[37] Mauchly's test of sphericity was significant for target race, $\chi^2(2) = 42.57$, $p < .001$; therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .66$).

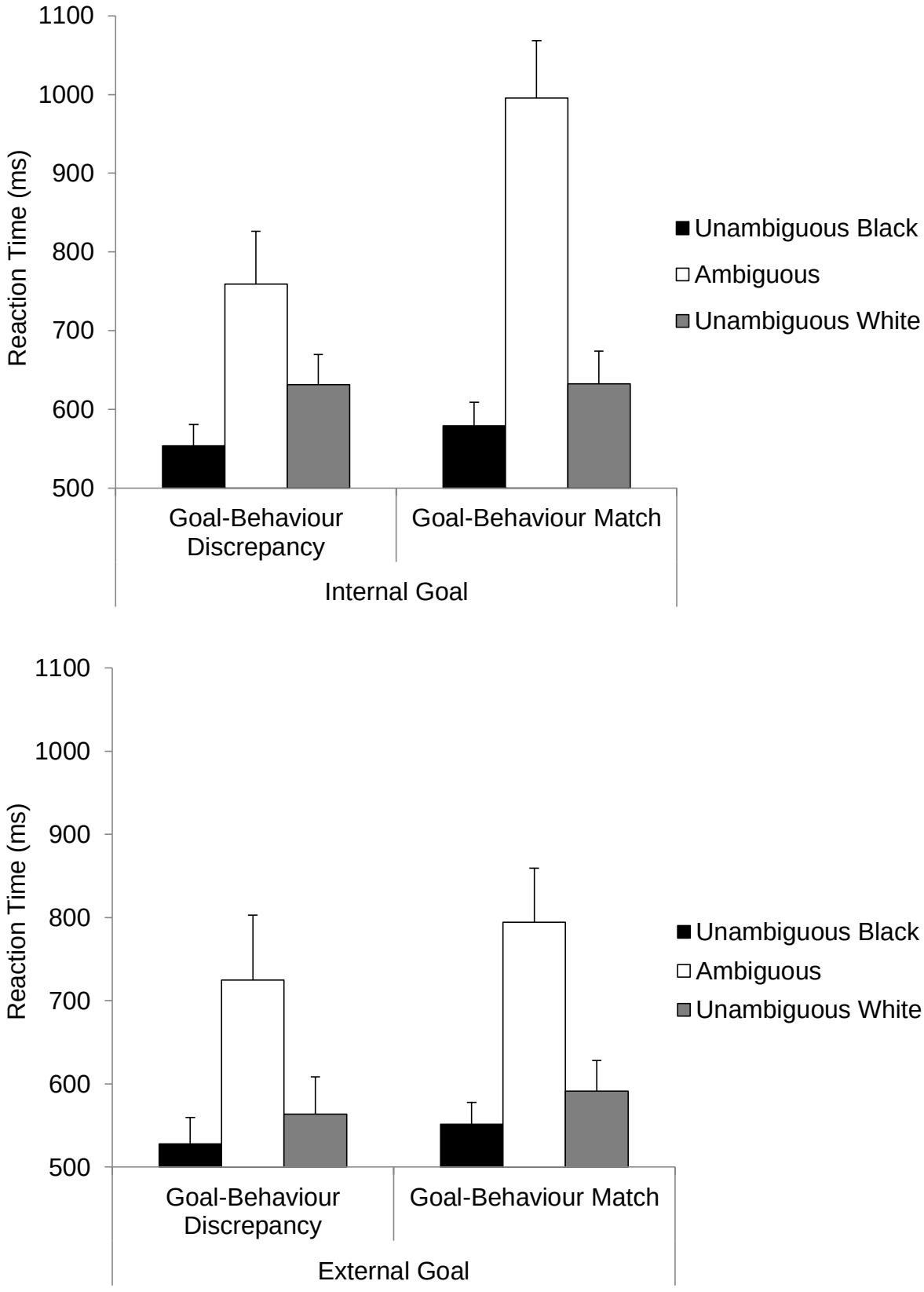

Figure 15. Mean reaction times (ms) as a function of goal source, goal–behaviour status, and target race, Experiment 6. *Note*. Error bars represent standard error.

2.2.3.1. Internal goal condition. The data were submitted to a 2 (Goal–Behaviour Status: goal–behaviour discrepancy vs. goal–behaviour match) × 3 (Target Race: unambiguous Black vs. racially ambiguous vs. unambiguous White) mixed-model ANOVA with goal–behaviour status as a between-participants factor. The analysis revealed a significant main effect of target race, $F(1.32, 30) = 40.58$, $p < .001$, $\eta^2_p = .57$[38]. Post-hoc tests using a Bonferroni correction demonstrated that participants categorised racially ambiguous faces (M = 877 ms, SE = 57.81) significantly slower than unambiguous Black faces (M = 567 ms, SE = 23.32; $p < .001$) and unambiguous White faces (M = 632 ms, SE = 31.65; $p < .001$); they were also faster to categorise unambiguous Black faces than White faces ($p = .007$).

More importantly, the analysis also revealed a significant Goal–Behaviour Status × Target Race interaction, $F(1.32, 30) = 6.30$, $p = .010$, $\eta^2_p = .17$[38]. Post hoc tests using a Bonferroni correction were conducted. In the goal–behaviour discrepancy condition, participants categorised racially ambiguous faces (M = 760 ms, SE = 77.96) slower than unambiguous Black faces (M = 554 ms, SE = 31.45; $p = .005$) but not slower than unambiguous White faces (M = 632 ms, SE = 42.68; $p = .079$); they were also faster to categorise racially unambiguous Black faces than White faces ($p = .018$). In the goal–behaviour match condition, participants categorised racially ambiguous faces (M = 995 ms, SE = 85.40) slower than both unambiguous Black faces (M = 580 ms, SE = 34.45; $p < .001$) and unambiguous White faces (M = 633 ms, SE = 46.75; $p < .001$), but responded equally fast to unambiguous Black and White faces ($p = .23$).

2.2.3.2 External goal condition. The data were submitted to a 2 (Goal–Behaviour Status: goal–behaviour discrepancy vs. goal–behaviour match) × 3 (Target Race:

[38] Mauchly's test of sphericity was significant for target race, $\chi^2(2) = 21.76$, $p < .001$; therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .66$).

unambiguous Black vs. racially ambiguous vs. unambiguous White) mixed-model ANOVA with goal–behaviour status as a between-participants factor. Only the main effect of target race was significant, $F(1.34, 29) = 50.30$, $p < .001$, $\eta^2_p = .63$[39]. Post-hoc tests using a Bonferroni correction demonstrated that participants were slower to categorise racially ambiguous faces (M = 759 ms, SE = 40.01) than both unambiguous White faces (M = 577 ms, SE = 24.44; $p < .001$) and unambiguous Black faces (M = 539 ms, SE = 16.46; $p < .001$); they were also faster to categorise unambiguous Black faces than unambiguous White faces (p = .025).

2.2.4 Categorisation. The second main dependent variable was the percentage of Black categorisations for racially ambiguous faces, calculated by dividing the number of Black categorisations by the number of racially ambiguous faces, and then multiplying by 100 (with higher scores indicating a higher tendency to categorise racially ambiguous faces as "Black"). The percentage of "Black" categorisations for racially ambiguous faces was submitted to a two-way ANOVA with goal source and goal–behaviour status as between-participants factors. No main effects or interactions were significant, all $p > .35$.

A series of one sample t-tests were conducted to compare participants' categorisations of ambiguous faces to chance (50%). For participants in the internal goal condition, those who received false feedback highlighting a potential goal–behaviour discrepancy were equally likely to categorise ambiguous faces as White ingroup members or Black outgroup members (M = 50%, SE = 7.52), $t(17) = 0.06$, $p = .96$, $d = 0.03$. In contrast, those who received false feedback highlighting a potential goal–behaviour match were marginally more likely to categorise racially ambiguous faces as White ingroup members rather than Black outgroup

[39] Mauchly's test of sphericity was significant for target race, $\chi^2(2) = 19.65$, $p < .001$; therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .67$).

members (M = 38%, SE = 6.12), $t(14) = 2.14$, $p = .069$, $d = 1.05$. For participants in the external goal condition, participants were equally likely to categorise racially ambiguous faces as White ingroup members or Black outgroup members regardless of whether they received feedback highlighting a potential goal–behaviour discrepancy (M = 46%, SE = 7.96), $t(12) = 0.56$, $p = .59$, $d = 0.32$, or a goal–behaviour match (M = 44%, SE = 5.60), $t(18) = 0.99$, $p = .33$, $d = 0.47$.

2.2.5 Goal commitment. Overall, goal commitment was high (M = 6.26, SD = 1.58). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for goal commitment was conducted. The analysis revealed a significant main effect of goal commitment, $F(1, 61) = 4.80$, $p = .032$, $\eta^2_p = .07$, indicating higher goal commitment in the internal goal condition (M = 6.68, SE = 0.27) compared to the external goal condition (M = 5.82, SE = 0.28). No additional main effects or interactions were significant, all $p > .52$.

3.0 Discussion

For participants in the *internal* goal condition, the discrepancy manipulation had an interesting effect on categorisation times. In the goal–behaviour match condition, these individuals categorised racially ambiguous faces more slowly than unambiguous faces (Black or White). However, this difference was smaller in the goal–behaviour discrepancy condition, where they still categorised racially ambiguous faces more slowly than unambiguous Black faces but now categorised racially ambiguous and unambiguous White faces at equal speed. In contrast, participants in the *external* goal condition categorised racially ambiguous faces more slowly than unambiguous faces (Black or White), regardless of whether false feedback highlighted the potential for a goal–behaviour discrepancy or match.

Slower reaction times to racially ambiguous faces relative to unambiguous faces have been interpreted as representing either an accuracy motivation (Blascovich et al., 1997) or a motivation to ensure the integrity of the ingroup (e.g., by excluding outgroup members; Leyens & Yzerbyt, 1992; Blascovich et al., 1997; see also Castano et al., 2002). Indeed, greater response times are traditionally associated with individuation-based (or, at least, deeper) processing of faces rather than category-based (or, at least, more superficial) processing of faces (Brewer, 1988; Fiske & Neuberg, 1990). That participants in the internal goal condition who were primed with a potential goal–behaviour match (who presumably perceived themselves as satisfying their own egalitarian goal) and participants in the external goal condition (who presumably had the goal of satisfying an egalitarian goal set by others) both responded more slowly to racially ambiguous faces suggests that they were motivated to process the faces more deeply. It seems unlikely, first, that these participants were motivated to defend the integrity of the ingroup (in that there was no evidence for an ingroup-overexclusion effect; Leyens & Yzerbyt, 1992). This leaves open the possibility of an accuracy motivation; indeed, this account fits with the responses of "successful" internal and "successful" and "unsuccessful" external goal participants, who were all slower to categorise racially ambiguous than unambiguous faces and were equally likely to categorise those ambiguous faces as White (ingroup) and Black (outgroup).

The pattern for "unsuccessful" internal goal participants (i.e., participants in the *internal* goal condition who received false feedback that highlighted the potential for a goal–behaviour *discrepancy*) categorised ingroup members (unambiguous White faces) and racially ambiguous faces equally fast quickly, suggesting an absence of motivation to process the ambiguous faces more deeply and perhaps a motivation to demonstrate egalitarianism by treating ingroup and ambiguous faces similarly. Furthermore, these participants demonstrated

a tendency toward categorising racially ambiguous faces more often as White (ingroup) than Black (outgroup). This pattern of responding is more akin to an ingroup overinclusion effect; "unsuccessful" internal-goal participants may have categorised ambiguous targets as ingroup members (i.e., White), even when they might not be, to reaffirm their inclusiveness and demonstrate their commitment to egalitarianism.

3.1Implications for Fishbach and Colleagues' Model

The findings of Experiment 6 are somewhat consistent with Fishbach and colleagues' model. Fishbach and colleagues' propose that people with internal goals adhere to their egalitarian goal following a large goal–behaviour discrepancy, but relax their goal pursuit following a small goal–behaviour discrepancy. Fishbach and colleagues' also propose that people with external goals adhere to the egalitarian goal following a small goal–behaviour discrepancy, but reject the goal following a large goal–behaviour discrepancy.

In Experiment 6, perceiving a goal–behaviour discrepancy (i.e., failure) in the context of a personally chosen (and thus certain) goal presumably motivated individuals to defend their commitment to the goal; devoting less attention (i.e., responding more quickly) to racially ambiguous targets and including ambiguous targets within the ingroup may be a way of signalling that race is "unimportant" and thus that one is truly egalitarian—consistent with Fishbach and colleagues' model. However, perceiving a goal–behaviour match (i.e., success) in the context of a personally chosen goal, or perceiving either a goal–behaviour discrepancy or match in the context of an externally imposed goal, also appeared to motivate individuals to defend their commitment to the goal; devoting more attention (i.e., responding more slowly) to racially ambiguous targets and categorising ambiguous faces randomly as White (ingroup) or Black (outgroup) may be a way of signalling that race is "important" but that one is still egalitarian—inconsistent with Fishbach and colleagues model. However, the speeded

categorisation task used in Experiment 6 does not assess automatic categorisation; hence, Fishbach and colleagues' model may only be applicable in an egalitarian context when assessing stereotype activation and automatic categorisation, rather than stereotype use and controlled categorisation.

In conclusion, Chapter 4 aimed to explore the effect of internal and external egalitarian goals on social categorisation following false feedback that induced either a goal–behaviour discrepancy or a goal–behaviour match. This experiment suggested that "successful" participants with an internal egalitarian goal, and participants with and external egalitarian goal, were motivated to be accurate in their categorisations of racially ambiguous targets. In contrast, "unsuccessful" participants with an internal egalitarian goal were motivated to demonstrate their inclusiveness. Of course, due to a sample size being on the small size, the effects found within the present experiment need to be replicated with a larger sample size. To the best of my knowledge, this is the first investigation to consider the roles of goal source and goal–behaviour discrepancy size in relation to social categorisation.

CHAPTER 5

GENERAL DISCUSSION

1.0 Background and Aims

This thesis aimed to shed some light on the mechanism through which goals influence stereotype activation. Specifically, this thesis aimed to examine the impact of both goal source and goal–behaviour discrepancy size. To achieve this aim, I applied Fishbach and colleagues' model of goal progress and motivation to stereotype activation. The model asserts that motivation and goal adherence are a function of the discrepancy between actual and desired goal states, and whether movement towards a goal is construed in terms of progress or commitment to the goal.

Past empirical research has focused largely on identifying the impact that goals exert on stereotype activation, identifying which goals foster stereotype activation and which goals inhibit stereotype activation (see Kunda & Spencer (2003) and Moskowitz (2010) for reviews). Less attention has focused on the mechanism through which goals influence stereotype activation and prejudice. Plant and Devine (1998) were one of the first to investigate *how* goals influence prejudice, distinguishing between two types of chronic motivation: internal motivation to act nonprejudiced (arising from personal beliefs and values) and external motivation to act nonprejudiced (arising from societal pressure to conform to a nonprejudiced norm). They also examined the potential interplay between motivation and goal–behaviour discrepancies in relation to affect. They found that large (vs. small) goal–behaviour discrepancies resulted in higher guilt for people high in internal motivation, and higher threatened affect for people high in external motivation. Plant and Devine's research was restricted to affective responses, although they did speculate about the impact of large goal–behaviour discrepancies on prejudice regulation. They expected a large goal–behaviour

discrepancy to motivate goal-consistent behaviour (i.e., reduced prejudicial responses) designed to reduce the size of the discrepancy for people high in internal motivation, but a lack of goal-consistent or perhaps even goal-inconsistent behaviour (i.e., prejudicial or increased prejudicial responses) for people high in external motivation.

Subsequent empirical research has examined the influence of self-regulatory success or failure on stereotype activation and prejudice (Fehr & Sassenberg, 2010; Mann & Kawakami, 2011; Moskowitz, & Li, 2011; Moskowitz & Stone, 2012). Self-regulatory failure may potentially induce a goal–behaviour discrepancy, and self-regulatory success may potentially induce a goal–behaviour match. This research has confirmed Plant and Devine's speculation about *internal* motivation by demonstrating that following a self-regulatory failure, internally motivated people exhibit goal–consistent behaviour (i.e., stereotype control and low prejudice) following a self-regulatory failure (e.g., Fehr & Sassenberg, 2010; Moskowitz & Li, 2011; Moskowitz & Stone, 2012) but a lack of goal-consistent behaviour (i.e., stereotype activation and prejudice) following self-regulatory success (e.g., Mann & Kawakami, 2011; Moskowitz & Li, 2011; Moskowitz & Stone, 2012).

Prior to my research, only one investigation had examined the impact of self-regulatory failure on prejudice for externally motivated people, and found no evidence for differential self-regulation following failure versus success (Fehr & Sassenberg, 2010). Investigating the impact of internal *and* external motivation, and goal–behaviour discrepancies, on stereotype activation is important. Although many people personally adopt an egalitarian goal because they value acting fairly and tolerantly of others (internally-generated egalitarian goal), many other people adopt an egalitarian goal to avoid social sanctions that arise from failing to comply with external pressure to act egalitarian (externally-imposed egalitarian goal; Plant & Devine, 1998). Yet the impact of external goals

on stereotype activation has received much less empirical attention than the impact of internal goals on stereotype activation. Understanding the conditions under which people pursue external goals is important because this may help reduce prejudice and stereotype activation/use in these individuals. Additionally, self-determination theory (Deci & Ryan, 1985; see also Ryan & Deci, 2000) and the self-concordance model (Sheldon & Elliot, 1999) indicate that external goals may gradually become internalised, eventually becoming internal goals. Therefore, understanding when people adhere to external goals may help the development of strategies designed to foster internal motivation.

2.0 Stereotype Accessibility

2.1 Internal Egalitarian Goals

The work in the present thesis found that the size of the goal–behaviour discrepancy is important for internal egalitarian goals. When internally motivated individuals perceive a lack of progress toward the egalitarian goal after detecting a large goal–behaviour discrepancy, they exhibit little to no accessibility of Black stereotypes, suggestive of stereotype control (Experiment 2). When these individuals perceive progress toward the egalitarian goal after reflecting on a past success, or after detecting either a small goal–behaviour discrepancy or a goal–behaviour match, they exhibit accessibility of Black stereotypes, suggestive of stereotype activation (Experiments 3–5).

The finding that participants with an internal egalitarian goal compensate for a large (but not a small) goal–behaviour discrepancy by controlling the accessibility of Black stereotypes is consistent with the goal-shielding model (see Kruglanski et al., 2002; Shah et al., 2002; Shah & Kruglanski, 2002). According to Kruglanski and colleagues (Kruglanski et al., 2002; Shah et al., 2002; Shah & Kruglanski, 2002), goal-shielding occurs when commitment to a goal is high, and facilitates the pursuit of a focal goal by "facilitating the

detection and processing of goal-relevant stimuli (as well as other goals)" (Moskowitz & Li, 2011, p. 3) and inhibiting competing goals. Moskowitz and colleagues (Moskowitz & Li, 2011; Moskowitz et al., 2011) applied the goal-shielding hypothesis to internal egalitarian goals and argued that an aversive tension state arises when there is a goal–behaviour discrepancy and that this state persists until a person compensates for the goal–behaviour discrepancy by responding consistently with the internal egalitarian goal. During the aversive tension state, the internal egalitarian goal is shielded for completion, and because stereotype activation and application are antithetical with the idea of egalitarianism, the goal of stereotyping is inhibited (Moskowitz & Li, 2011; Moskowitz et al., 2011). However, once a person responds consistently with the internal egalitarian goal (by compensating for the large goal–behaviour discrepancy, or by perceiving a small goal–behaviour discrepancy), they affirm their egalitarian identity, shutting down the tension-state and goal-relevant responding (Moskowitz & Li, 2011; Moskowitz et al., 2011).

2.2 External Egalitarian Goals

The work in the present thesis also suggests that the size of the goal–behaviour discrepancy *may not* be important for external egalitarian goals. When externally motivated individuals perceive a lack of progress toward the egalitarian goal after detecting a large goal–behaviour discrepancy (Experiment 2), they exhibit accessibility of Black stereotypes (suggestive of stereotype activation). Similarly, when these individuals perceive progress toward the egalitarian goal after detecting either a small goal–behaviour discrepancy or a goal–behaviour match (Experiments 3–5), they also exhibit accessibility of Black stereotypes (suggestive of stereotype activation).

The finding that participants with an external egalitarian goal fail to respond consistently with the goal following a small (but not a large) goal–behaviour discrepancy by

controlling the accessibility of Black stereotypes is inconsistent with the goal-shielding model (see Kruglanski et al., 2002; Shah et al., 2002; Shah & Kruglanski, 2002). According to Fishbach and colleagues' model, perceiving a small goal–behaviour discrepancy in relation to an external goal signals that commitment to the goal is high. Kruglanski and colleagues (Kruglanski et al., 2002; Shah et al., 2002; Shah & Kruglanski, 2002) argue that goal-shielding occurs when commitment to a goal is high, and facilitates the pursuit of a focal goal by "facilitating the detection and processing of goal-relevant stimuli (as well as other goals)" (Moskowitz & Li, 2011, p. 3) and inhibiting competing goals. Indeed, past research by Fishbach and Dhar (2005) found that when commitment to a higher-order academic goal was uncertain (and uncertain commitment should characterise external goals), successfully making progress toward the academic goal by studying resulted in greater interest in academic activities compared to competing social activities, presumably because the uncertain commitment was boosted.

However, it is important to note that the findings of Experiment 1 demonstrated that when individuals with an external egalitarian goal perceive progress toward the egalitarian goal after detecting a small goal–behaviour discrepancy, they exhibit little to no accessibility of Black stereotypes (suggestive of control over stereotype activation). But why do individuals with an external egalitarian goal exhibit little to no accessibility of negative Black stereotypes after a small goal–behaviour discrepancy in Experiment 1, but accessibility of negative Black stereotypes after a small goal–behaviour discrepancy or goal–behaviour match in Experiments 3–5? The difference in findings across experiments may be due to a change in discrepancy manipulation. In Experiment 1, goal progress was manipulated through self-reflection on a past experience. In Experiments 3–5, however, discrepancies were manipulated by giving

participants an ostensible test of egalitarian behaviour (a modified IAT) followed by external feedback about their performance.

This change in manipulation could account for the difference in findings. First, recalling a past successful egalitarian act may make participants feel more successful than receiving feedback on the basis of a "test". Research has demonstrated that participants have high expectations about their own performance on the IAT (Howell et al., 2013), which formed the basis of the "test" in Experiments 3–5 of the present thesis. For example, Howell et al. (2013) demonstrated that the majority of participants expected to favour Black and White people equally (62%) or to exhibit a slight automatic preference for White individuals (21%)[40]. Therefore, exceeding participants expectations on the basis of a test might be simply difficult to do. Indeed, Sassenberg and colleagues (2011, Experiment 2) demonstrated that participants' responded neutrally on affect items when they received false feedback indicating that their studying behaviour matched other members of their ingroup because their expectations were not exceeded. In Experiments 3–5 of the present thesis, receiving feedback at 98% or 100% may have matched participants' expectations: Indeed participants felt only moderately to very positive following feedback at 98% and 100%, and in Experiment 6 participants indicated their the feedback at 100% matched their expectations very well. However, when participants described a time in the past where they successfully acted egalitarian towards a Black person, as in Experiment 1, multiple instances of past successful egalitarian acts might actually come to mind before one example is selected. Consequently, participants may feel like they are egalitarian all of the time and have achieved the egalitarian goal. However, this is purely speculation: Participants in Experiment 1 did not indicate how

[40] Of the remaining 17%, 3% expected to demonstrate a "strong automatic preference for Black individuals", 8% expected to demonstrate a "slight automatic preference for Black individuals", and 6% expected to demonstrate a moderate automatic preference for Black individuals (Howell et al., 2013, p. 3).

successful they felt after reflecting on a past success, so I cannot be sure that they felt more successful after writing about a past success than after receiving positive feedback following the fake "test". In future research, when using an ostensible assessment of participants' egalitarian behaviour, it might be necessary to provide negative feedback initially (to counter participants' initially high expectations) and then increase to a small goal–behaviour discrepancy size (see Mann & Kawakami, 2011, for an example of this manipulation with an internal nonprejudiced goal).

Second, when participants learned that they would be completing a test of how motivated they are to act egalitarian towards Black people, they may have felt pressured to control their responses in order to appear nonprejudiced (see Hausmann & Ryan, 2004). For people in the internal goal condition, pressure to appear egalitarian on the test is congruent with their personal egalitarian beliefs (see Gordijn et al., 2004). As a result, these individuals are less likely to have expended effort controlling their responses during the test, leaving ample cognitive resources available to control stereotype activation (see Gordijn et al., 2004). For people in the external goal condition, however, pressure to appear egalitarian on the test is incongruent with their personal nonegalitarian beliefs (see Gordijn et al., 2004). As a result, these individuals may have needed to exert a great deal of effort to control their responses during the test because the egalitarian and nonegalitarian response options competed with each other, depleting their cognitive resources (see Gordijn et al., 2004; Hausmann & Ryan, 2004). As the restoration of cognitive resources takes time, and stereotype suppression requires ample cognitive resources (Macrae et al., 1994; Wyer et al., 2000), these individuals may have exhibited stereotype rebound in my Experiments 2–5. In contrast, in Experiment 1 it seems less likely that describing a past successful egalitarian act towards a Black person that arose from social or political pressure would have required the suppression of stereotypes

or prejudice. Consequently, these individuals may have had ample cognitive resources available to control stereotype activation. Future research should determine whether goal–behaviour discrepancy size matters for people with external egalitarian goals by using a manipulation of goal–behaviour discrepancy size that does not deplete cognitive resources and increase the likelihood of stereotype rebound as in the work in the present thesis and in past research (e.g., Fehr & Sassenberg, 2010).

2.3 Conclusion: Stereotype Activation

This thesis aimed to shed some light on the mechanism through which goals influence stereotype activation. The findings of the present thesis suggest that the source of the goal (i.e., internally-generated vs. externally-imposed) is an important part of the mechanism through which goals influence stereotype activation. Additionally, discrepancy size also appears to be an important part of the mechanism through which internal, but not external, goals influence stereotype activation. However, I proposed that it is plausible that (1) making participants feel successful enough on the basis of a "test" is simply difficult to do, and (2) the manipulation of goal–behaviour status may have cognitively taxed participants in the external goal condition. Consequently, future research should confirm whether discrepancy size does or does not matter for external egalitarian goals.

3.0 Categorisation

This thesis also aimed to determine when egalitarian goals take effect. To identify the mechanism through which internal and external egalitarian goals influence stereotype activation, it is important to understand whether higher-order goals like the goal to act egalitarian influence social categorisation. In Experiment 6, I examined the effect of goal–behaviour discrepancies on social categorisation for individuals with either an internal or external egalitarian goal.

3.1 Internal Egalitarian Goals

Experiment 6 found that the size of the goal–behaviour discrepancy is important for internal egalitarian goals. When these individuals perceive progress toward the egalitarian goal after detecting a goal–behaviour discrepancy match, they appear to demonstrate their egalitarian attitude by responding accurately when categorising racially ambiguous faces. First, these individuals responded more slowly to racially ambiguous faces suggesting that they were motivated to process the faces more deeply. Second, these individuals were equally likely to categorise those unambiguous faces as White (ingroup) and Black (outgroup). Overall, this pattern of responding is consistent with a multiculturalism ideology which "advocates considering, and sometimes emphasising and celebrating, category memberships" (Richeson & Nussbaum, 2004, p. 417). If participants had the opportunity to categorise faces as "White", "Black", or "multiracial", like internally motivated participants in Chen and colleagues (Chen et al., 2014) research, they may have utilised all three categories in order to celebrate different category memberships.

When these individuals perceive a lack of progress after detecting a goal–behaviour discrepancy, they appear to demonstrate their egalitarian attitude by treating ingroup members (unambiguous White faces) and racially ambiguous faces similarly. First, these individuals categorised ingroup members (unambiguous White faces) and racially ambiguous faces equally fast quickly, suggesting an absence of motivation to process the ambiguous faces more deeply. Second, these individuals were concerned about being inclusive; hence, these individuals were more likely to categorise ambiguous targets as ingroup members (i.e., White), even when they might not be, to reaffirm their inclusiveness and demonstrate their commitment to egalitarianism. Overall, this pattern of responding is consistent with a colour-blind ideology which "advocate[s] reducing, eliminating, and ignoring category

memberships" (Richeson & Naussbaum, 2004, p. 417). Indeed, these individuals may have recategorised racially ambiguous faces into the ingroup (see Gaertner, Dovidio, Anastasio, Bachman, & Rust, 1993), and consequently relied more on category-based processing (hence faster reaction times) by identifying facial features that are diagnostic of their own racial group in the racially ambiguous faces and including them within the ingroup.

3.2 External Egalitarian Goals

Experiment 6 found that the size of the goal–behaviour discrepancy is *not* important for external egalitarian goals. When these individuals perceive progress toward the egalitarian goal after detecting a goal–behaviour match, or a lack of progress toward the egalitarian goal after detecting a goal–behaviour discrepancy, they appear to demonstrate their egalitarian attitude by responding accurately when categorising racially ambiguous faces. First, these individuals responded more slowly to racially ambiguous faces suggesting that they were motivated to process the faces more deeply. Second, these individuals were equally likely to categorise those unambiguous faces as White (ingroup) and Black (outgroup). Overall, this pattern of responding is identical to that exhibited by participants in the internal goal condition who received false feedback that highlighted the potential for a goal–behaviour match. Instead of responding consistently with a multicultural ideology like "successful" participants in the internal goal condition, individuals in the external goal condition may have been responding accurately in order to avoid mistakenly categorising a racially ambiguous face as "Black", which could potentially be construed as biased responding (i.e., only prototypical White faces are "White". This is a tentative interpretation given that the present experiment does not allow me to determine conclusively whether the accuracy motivation arises due to an accuracy goal or a goal to avoid bias, but is based on research by Chen and colleagues (Chen et al., 2014). They found that people high in external motivation are

cautious about using “multiracial”, even when racial ambiguity is high, perhaps because they fear the term is not politically correct and therefore socially unacceptable. In our experiment, when participants with an external goal were forced to choose between monoracial categories (i.e., White or Black), they categorised ambiguous faces slower and randomly, perhaps because they feared responding in a socially unacceptable manner.

3.4 Conclusion: Categorisation

This thesis aimed to determine when egalitarian goals take effect. To identify the mechanism through which internal and external egalitarian goals influence stereotype activation, it is important to understand whether higher order goals like the goal to act egalitarian influence social categorisation. The findings of the present thesis suggest that the source of the goal (i.e., internal vs. external) is an important part of the mechanism through which goals influence social categorisation. Additionally, discrepancy size also appears to be an important part of the mechanism through which *internal,* but not external, goals influence social categorisation. Additionally, the results suggest, albeit tentatively, that beliefs about what goal success looks like differ as a function of goal source. Individuals with an internal egalitarian goal who perceived a goal–behaviour discrepancy appear to believe that goal success means treating ingroup members (unambiguous White faces) and racially ambiguous faces as the same (i.e., being inclusive). In contrast, individuals with an external goal regardless of goal–behaviour discrepancy size, and individuals with an internal goal who perceived a goal–behaviour match, appear to believe that goal success means accurately categorising faces.

4.0 Additional Future Directions

4.1 Types of Internal and External Goals

According to Plant and Devine (1998), people can be motivated to act nonprejudiced primarily for internal reasons, primarily for external reasons, or for both internal and external reasons. In the present thesis, I examined the influence of temporary internal versus external egalitarian goal son stereotype accessibility. Future research should examine the influence of a temporary *combined* internal/external egalitarian goal on stereotype accessibility. According to Plant and Devine (1998, 2009) when people are both internally and externally motivated, internal motivation is the primary motivation. In line with this conclusion, Plant and Devine (2009) demonstrated that while individuals high in both internal and external motivation had both the intention to eradicate prejudice altogether (also held by individuals high in internal but low in external motivation) and the intention to hide prejudice from others (also held by individuals low in internal but high in external motivation), the intention to eradicate prejudice altogether was primary: When given the opportunity to use a program that would reduce detectable prejudice in the short-term but inadvertently increase undetectable prejudice in the long-term, these individuals were not interested. The research by Plant and Devine (2009) suggests that if a person holds an internal and an external goal, we should expect them to ask questions about their progress towards the goal and demonstrate the same self-regulatory pattern for internal goals within Fishbach and colleagues' model.

Additionally, future research should also examine whether the source of the external goal (e.g., peer versus expert) changes the influence of the external egalitarian goal on stereotype accessibility. To the best of my knowledge, no research has examined this issue. It is possible to speculate that an external goal imposed by a peer might be more effective than an external goal imposed by an expert. This speculation is based on past research, which has

found that individuals who identify strongly with a group "take on" the group's goals, and consequently view the group's goals as internally generated rather than externally imposed. Given the higher likelihood that people would be more likely to identify with a peer than an expert, and should thus be more likely to adopt the external goal as their own internally generated goal, using peers to impose external goals may be a more effective means of transforming an external goal into an internal goal than using experts.

4.2 Refinement of Measures

4.2.1 Identification. Using an external reference group that is smaller in size and higher in entitativity than the White ethnic group or the egalitarian group would be useful. For example, using University of Birmingham students as the external reference group would provide a better assessment of whether identification moderates the influence of goal source and goal–behaviour discrepancies on stereotype accessibility. Empirical research by Fishbach and colleagues' and Sassenberg et al. (2011) has both used University students as the external reference group, which may be the reason why these investigations have found that identification has a moderating role. In a follow-up study, it would be beneficial to determine whether participants who highly identify with fellow University students adopt the egalitarian goal as their own personally-generated goal.

4.2.2 Competing goal. As positive distinctiveness is incompatible with acting egalitarian, I explored whether positive distinctiveness serves as a competing goal that internally motivated people seek when they perceive a small goal–behaviour discrepancy. Specifically, I predicted that participants in the internal goal condition who held a competing positive distinctiveness goal would exhibit greater stereotype accessibility than those who did not hold a competing positive distinctiveness goal because the increased stereotype accessibility would aid the competing positive distinctiveness goal. In contrast to my

predictions, positive distinctiveness did not moderate the influence of goal source on stereotype accessibility in Experiments 2–5 of the present thesis. However, given the existence of a pervasive egalitarian social norm and the general desire to maintain an egalitarian self-image (see Fehr & Sassenberg, 2010; Fehr et al., 2012; Gaertner & Dovidio, 1986; McConahay, 1986; Monteith et al., 1996), it may be beneficial to use an implicit measure of positive distinctiveness rather than the explicit measure of positive distinctiveness used in Experiments 1–5 to avoid social desirable responding. Social identity theory (Tajfel & Turner, 1979, 1986) argues that ingroup favouritism is the concrete manifestation of positive distinctiveness. An implicit measure of ingroup favouritism may moderate the influence of goal source on stereotype accessibility. Indeed, Mann and Kawakami (2011) demonstrated that internally motivated participants sat closer to a White confederate after successfully progressing (vs. failing to progress) towards the goal to evaluate Black faces positively.

4.2.3 Goal commitment. Based on Fishbach and colleagues' model, it might be expected that goal commitment would be higher in the internal than the external goal condition. Yet self-reported goal commitment did not differ between the goal conditions throughout Experiments 2–5. In Experiments 2–5, I measured participants' commitment to a temporary egalitarian goal 5–10 minutes after assessing their motivation to pursue the egalitarian goal (assessing stereotype activation/control). Furthermore, my assessment of participants' motivation to control their stereotypes allowed participants to reduce the discrepancy and satisfy their motivation to pursue the egalitarian goal. In contrast, Koo and Fishbach (2008) measured participants' commitment to a temporary academic goal (studying for an exam) immediately after assessing their motivation to pursue the academic goal (hours they intend to spend studying and the amount of effort they intent to expend). Furthermore, Koo and Fishbach's assessment of motivation did not allow participants to reduce the

discrepancy and satisfy their motivation to pursue the academic goal. It is feasible that *temporary* internal and external goals may *temporarily* boost or undermine goal commitment, such that once the goal–behaviour discrepancy is reduced, goal commitment returns to the level prior to the goal manipulation. If I had examined chronic motivation instead of temporary goals, I may have found that goal commitment is lower for chronic external motivation than chronic internal motivation no matter when goal commitment is assessed.

5.0 General Conclusion

Empirical research has focused largely on identifying the impact that goals exert on stereotyping. Less attention has focused on the mechanism through which goals influence stereotype activation and prejudice. This thesis aimed to shed some light on the mechanism through which internally-generated and externally-imposed goals influence stereotype activation. The work presented in this thesis demonstrated that the source of the goal (i.e., internally-generated vs. externally-imposed) is an important part of the mechanism through which goals influence social categorisation and stereotype activation. Additionally, discrepancy size also appears to be an important part of the mechanism through which *internal,* but not external, goals influence social categorisation and stereotype activation.

REFERENCES

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

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology, 91,* 652–661. doi:10.1037/0022-3514.91.4.652

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology, 84,* 738–753. doi:10.1037/0022-3514.84.4.738

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neutral signals for control. *Journal of Personality and Social Psychology, 94,* 60–74. doi:10.1037/0022-3514.94.1.60

Ashmore, R. D. (1970). Prejudice: Causes and cures. In B. E. Collins (Ed.), *Social psychology: Social influence, attitude change group processes, and prejudice* (pp. 245–339). Reading, MA: Addison Wesley.

Badea, C., Jetten, J., Czukor, G., & Askevis-Leherpeux, F. (2010). The bases of identification: When optimal distinctiveness needs face social identity threat. *British Journal of Social Psychology, 49,* 21–41. doi:10.1348/000712608X397665

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology, 87,* 5–22. doi:10.1037/0022-3514.87.1.5

Bargh, J. A. (1990). Auto-motives: Preconscious determinants of social interaction. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 93–130). New York, NY: Guilford.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 361–382). New York, NY: Guilford.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54,* 462–479. doi:10.1037//0003-066X.54.7.462

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. Reis & C. Judd (Eds.), *Research methods in social psychology* (pp. 253–285). New York: Cambridge University Press.

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

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

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology, 70,* 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,* 828–841. doi: I0.1037//0022-3514.81.5.828

Blanchard, F. A., Lilly, T., & Vaughn, L. A. (1991). Reducing the expression of racial prejudice. *Psychological Science, 2,* 101–105. doi:10.1111/j.1467-9280.1991.tb00108.x

Blascovich, J., Wyer, N. A., Swart, L. A., & Kibler, J. L. (1997). Racism and racial categorization. *Journal of Personality and Social Psychology, 72,* 1364–1372. doi:10.1037//0022-3514.72.6.1364

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer Jr. (Ed.), *Advances in social cognition* (Vol. 11, pp. 1–52). Mahwah, NJ: Erlbaum.

Bodenhausen, G. V., Todd, A. R., & Richeson, J. (2009). Controlling prejudice and stereotyping: Antecedents, mechanisms, and contexts. In T. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 111–135). New York: Psychology Press.

Bonezzi, A., Brendl, C. M., & De Angelis, M. (2011). Stuck in the middle: The psychophysics of goal pursuit. *Psychological Science, 22,* 607–612. doi:10.1177/0956797611404899

Bradley-Geist, J. C., King, E. B., Skorinko, J., Hebl, M. R., & McKenna, C. (2010). Moral credentialing by association: The importance of choice and relationship closeness. *Personality and Social Psychology Bulletin, 36,* 1564–1575. doi:10.1177/0146167210385920

Brehm, J. W. (1966). *A theory of psychological reactance*. Oxford, UK: Academic Press.

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

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin, 17,* 475–482. doi:10.1177/0146167291175001

Butz, D. A., & Plant, E. A. (2009). Prejudice control and interracial relations: The role of motivation to respond without prejudice. *Journal of Personality, 77,* 1311–1342. doi:10.1111/j.1467-6494.2009.00583.x

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science, 3,* 14–25. doi:10.1002/bs.3830030103

Castano, E., Yzerbyt, V., Bourguignon, D., & Seron, E. (2002). Who may enter? The impact of in-group identification on in-group/out-group categorization. *Journal of Experimental Social Psychology, 38,* 315–322. doi:10.1006/jesp.2001.1512

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

Chen, J. M., Moons, W. G., Gaither, S. E., Hamilton, D. L., & Sherman, J. Q. (2014). Motivation to control prejudice predicts categorization of multiracials. *Personality and Social Psychology Bulletin*. Advance online publication. doi:10.1177/0146167213520457

Crisp, R. J., & Beck, S. R. (2005). Reducing intergroup bias: The moderating role of ingroup identification. *Group Processes and Intergroup Relations, 8,* 173–185. doi:10.1177/1368430205051066

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

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24,* 349–354. doi: 10.1037/h0047358

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12,* 163–170. doi:10.1111/1467-9280.00328

Dasgupta, N. & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81,* 800–814. doi: I0.1037//0022-3514.81.5.800

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York, NY: Plenum Press.

DeCoster, J. (2012). *Converting effect sizes.* Retrieved from: http://www.stat-help.com/spreadsheets.html

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18. doi:10.1037//0022-3514.56.1.5

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin, 21,* 1139–1150. doi:10.1177/01461672952111002

Devine, P. G., & Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology, 60,* 817–830. doi:10.1037//0022-3514.60.6.817

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology, 82,* 835–848. doi:10.1037//0022-3514.82.5.835

Dovidio, J. F., & Gaertner, S. L. (1986). Prejudice, discrimination, and racism: Historical trends and contemporary approaches. In J. F. Dovidio & S. L. Gartner (Ed.), *Prejudice, discrimination, and racism* (pp. 1–30). Orlando, FL: Academic Press.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33,* 510–540. doi:10.1006/jesp.1997.1331

Dowd, E. T., Milne, C. R., & Wise, S. L. (1991). The therapeutic reactance scale: A measure of psychological reactance. *Journal of Counseling and Development, 69,* 541–545. doi:10.1002/j.1556-6676.1991.tb02638.x

Dutton, D. G., & Lake, R.A. (1973). Threat of own prejudice and reverse discrimination in interracial situations. *Journal of Personality and Social Psychology, 28,* 94–100. doi:10.1037/h0035582

Effron, D. A., Cameron, J. S., & Monin, B. (2009). Endorsing Obama licenses favoring Whites. *Journal of Experimental Social Psychology, 45,* 590–593. doi:10.1016/j.jesp.2009.02.001

Effron, D. A., Miller, D.T., & Monin, B. (2012). Inventing racist roads not taken: The licensing effect of immoral counterfactual behaviors. *Journal of Personality and Social Psychology, 103,* 916–932. doi:10.1037/a0030008

Empirisoft Corporation. (2004). MediaLab and DirectRT v2004 [Computer software]. New York, NY: Author.

Equality and Human Rights Commission. (2009). *Police and racism: What has been achieved 10 years after the Stephen Lawrence Inquiry Report?* Retrieved from: http://www.equalityhumanrights.com/uploaded_files/raceinbritain/policeandracism.pdf

Esses, V. M., Haddock, G., & Zanna, M. P. (1993). Values, stereotypes, and emotions as determinants of intergroup attitudes. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in group perception* (pp. 137–166). San Diego, CA: Academic Press.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149–1160. doi:10.3758/BRM.41.4.1149

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69,* 1013–1027. doi:10.1037/0022-3514.69.6.1013

Fehr, J., & Sassenberg, K. (2010). Willing and able: How internal motivation and failure help to overcome prejudice. *Group Processes & Intergroup Relations, 13,* 167–181. doi:10.1177/1368430209343116

Fehr, J., Sassenberg, K., & Jonas, K. J. (2012). Willful stereotype control: The impact of internal motivation to respond without prejudice on the regulation of activated stereotypes. *Journal of Psychology, 220,* 180–186. doi:10.1027/2151-2604/a000111

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Tests. *Basic and Applied Social Psychology, 27,* 307–316. doi:10.1207/s15324834basp2704_3

Fishbach, A. (2009). The dynamics of self-regulation. In J. R., Forgas , R. F., Baumeister, & D. M., Tice (Eds.), *Psychology of Self-Regulation: Cognitive, Affective and Motivational Processes*, (pp. 163-181). New York: Psychology Press

Fishbach, A., & Dhar, R. (2005). Goals as excuses or guides: The liberating effect of perceived goal progress on choice. *Journal of Consumer Research, 32,* 370–377. doi:10.1086/497548

Fishbach, A., & Dhar, R. (2007). Dynamics of goal-based choice: Toward an understanding of how goals commit versus liberate choice. In C. P. Haugtvedt, P. M. Herr & F. R. Kardes (Eds.), *Handbook of consumer psychology* (pp. 611–637). New York, NY: Psychology Press.

Fishbach, A., Henderson, M. D., & Koo, M. (2011). Pursuing goals with others: Group identification and motivation resulting from things done versus things left undone. *Journal of Experimental Psychology: General, 140,* 520–534. doi:10.1037/a0023907

Fishbach, A., Zhang, Y., & Koo. (2009). The dynamics of self-regulation. *European Review of Social Psychology, 20,* 315–344. doi:10.1080/10463280903275375

Fishbach, A., Zhang, Y., & Koo. (2006). Subgoals as substitutes or complements: The role of goal accessibility. *Journal of Personality and SocialPsychology, 91,* 232–242. doi:10.1037/0022-3514.91.2.232

Fiske, S. T. & Neuberg, S. L. (1990). A continuum model of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). San Diego, CA: Academic Press.

Fiske, S. T., Lin, M., and Neuberg, S. L. (1999). The continuum model ten years later. In S. Chaiken and Y. Troppe (Eds.), *Dual-process theories in social psychology* (pp. 231–254). New York: Guilford Press.

Förster, J., Grant, H., Idson, L. C., & Higgins, E. T. (2001). Success/failure feedback, expectancies, and approach/avoidance motivation: How regulatory focus moderates classic relations. *Journal of Experimental Social Psychology, 37,* 253–260. doi:10.1006/jesp.2000.1455a

Förster, J., Higgins, E. T., & Idson, L. C. (1998). Approach and avoidance strength during goal attainment: Regulatory focus and the "goal looms larger" effect. *Journal of Personality and Social Psychology, 75,* 1115–1131. doi:10.1037/0022-3514.75.5.1115

Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). New York, NY: Academic Press.

Gaertner, S.L., Dovidio, J.F., Anastasio, P.A., Bachman, B.A., & Rust, M.C. (1993). The common ingroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology, 4,* 1-26.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology, 78,* 708–724. doi:10.1037//0022-3514.78.4.708

Galinsky, A. D., & Moskowitz, G. B. (2007). Further ironies of suppression: Stereotype and counterstereotype accessibility. *Journal of Experimental Social Psychology, 43,* 833–841. doi:10.1016/j.jesp.2006.09.001

Garvey, A. M. (2011). *Sticking to it? How consumer goal progress affects goal perseverance*. (Unpublished doctoral dissertation). Pennsylvania State University, USA.

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

Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B., & Strack, F. (2008). When "Just Say No" is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology, 44,* 370–377. doi:10.1016/j.jesp.2006.12.004

Gawronski, B., Peters, K. R., Brochu, P. M., & Strack, F. (2008). Understanding the relations between different forms of racial prejudice: A cognitive consistency perspective. *Personality and Social Psychological Bulletin, 34,* 648–665. doi:10.1177/0146167207313729

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

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *The American Psychologist, 54,*493–503. doi:10.1037//0003-066X.54.7.493

Gordijn, E. H., Hindriks, I., Koomen, W., Dijksterhuis, A., & Van Knippenberg, A. (2004). Consequences of stereotype suppression and internal suppression motivation: A self-regulation approach. *Personality and Social Psychology Bulletin, 30,* 212–224. doi:10.1177/0146167203259935

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

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

Hamilton, D. L., & Trolier, T. K. (1986). Stereotypes and stereotyping: An overview of the cognitive approach. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 127–163). Orlando, FL: Academic Press.

Hausmann, L. R. M., & Ryan, C. S. (2004). Effects of external versus internal motivation to control prejudice on implicit prejudice: The mediating role of efforts to control prejudiced responses. *Basic and Applied Social Psychology, 26,* 215–225. doi:10.1080/01973533.2004.9646406

Henry, P. J., & Sears, D. O. (2002). The symbolic racism 2000 scale. *Political Psychology, 23,* 253–283. doi:10.1111/0162-895X.00281

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94,* 319–340. doi:10.1037//0033-295X.94.3.319

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52,* 1280–1300.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). New York: Academic Press

Hogg, M. A., Sherman, D. K., Dierselhuis, J., Maitner, A.T., & Moffitt, G. (2007). Uncertainty, entitativity, and group identification. *Journal of Experimental Social Psychology, 43,* 135–142. doi:10.1016/j.jesp.2005.12.008

Howell, J. L., Collisson, B., Crysel, L., Garrido, C. O., Newell, S. M., Cottrell, C. A., Smith, C. T., & Shepperd, J. A. (2013). Managing the threat of impending implicit attitude feedback. *Social Psychological and Personality Science, 4,* 714–720. doi:10.1177/1948550613479803

Hull, C. L. (1932). The goal-gradient hypothesis and maze learning. *Psychological Review, 39,* 25–43. doi: 10.1037/h0072640

Johns, M., Cullum, J., Smith, T., & Freng, S. (2008). Internal motivation to respond without prejudice and automatic egalitarian goal activation. *Journal of Experimental Social Psychology, 44,* 1514–1519. doi:10.1016/j.jesp.2008.07.003

Kawakami, K., Dovidio, J. F., Mill, 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,* 871–888. doi:10.1037//0022-3514.78.5.871

Kawakami, K., Dovidio, J. F., & van Kamp, S. (2007). The impact of counterstereotypic training and related correction processes on the application of stereotypes. *Group Processes & Intergroup Relations, 10,* 139–156. doi:10.1177/1368430207074725

Kawakami, K., Phills, C. E., Steele, J. R., & Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder: Improving implicit racial attitudes and interracial interactions through approach behaviours. *Journal of Personality and Social Psychology, 92,* 957–971. doi:10.1037/0022-3514.92.6.957

Klonis, S. C., Plant, E. A., & Devine, P. G. (2005). Internal and external motivation to respond without sexism. *Personality and Social Psychology Bulletin, 31,* 1237–1249. doi:10.1177/0146167205275304

Kivetz, R., Urminsky, O., & Zheng, Y. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *Journal of Marketing Research, 43,* 39–58. doi:10.1509/jmkr.43.1.39

Koo, M., & Fishbach, A. (2008). Dynamics of self-regulation: How (un)accomplished goal actions affect motivation. *Journal of Personality and Social Psychology, 94,* 183–195. doi:10.1037/0022-3514.94.2.183

Kruglanski, A. W., Shah, J. Y., Fishbach, A., Friedman, R., Chun, W. Y., & Sleeth-Keppler, D. (2002). A theory of goal systems. *Advances in Experimental Social Psychology, 34,* 331-378. doi: 10.1016/S0065-2601(02)80008-9

Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin, 129,* 522–544. doi:10.1037/0033-2909.129.4.522

Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science, 22,* 1472–1477. doi:10.1177/0956797611427918

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72,* 275–287. doi:10.1037/0022-3514.72.2.275

Leyens, J-Ph., & Yzerbyt, V. Y. (1992). The ingroup overexclusion effect: Impact of valence and confirmation on stereotypical information search. *European Journal of Social Psychology, 22,* 549–569.

Liberman, N., & Förster, J. (2000). Expression after suppression: A motivational explanation of postsuppressional rebound. *Journal of Personality and Social Psychology, 79,* 190–203. doi:10.1037//0022-3514.79.2.190

Lippman, W. (1922). *Public opinion.* New York, NY: Harcourt & Brace.

Livingston, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology, 82,* 5–18. doi:10.1037//0022-3514.82.1.5

Locke, E., & Latham, G. (2005). Goal-setting theory. In J. B. Miner (Ed.), *Organizational behaviour 1: Essential theories of motivation and leadership* (pp. 159–199). London, UK: M.E. Sharpe.

Loo, R., & Loewen, P. (2004). Confirmatory factor analyses of scores from full and short versions of the Marlowe-Crowne social desirability scale. *Journal of Applied Social Psychology, 34,* 2343–2352. doi:10.1111/j.1559-1816.2004.tb01980.x

Loo, R., & Thorpe, K. (2000). Confirmatory factor analyses of the full and short versions of the Marlowe-Crowne social desirability scale. *The Journal of Social Psychology, 140,* 628–635. doi:10.1080/00224540009600503

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81,* 842–855. doi:10.1037//0022-3514.81.5.842

Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology, 69,* 397–407. doi:0022-3514.69.3.397

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

Macrae, C. N., Bodenhausen, G. V., Milne, A. D., Thorn, T. M. J., & Castelli, L. (1997). On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology, 33,* 471–489. doi:10.1006/jesp.1997.1328

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

Macrae, C. N., Quinn, K. A., Mason, M. F., & Quadflieg, S. (2005). Understanding others: The face and person construal. *Journal of Personality and Social Psychology, 89,* 686–695. doi:10.1037/0022-3514.89.5.686

Macrae, C. N., Stangor, C., & Milne, A. B. (1994). Activating social stereotypes: A functional analysis. *Journal of Experimental Social Psychology, 30,* 370–389. doi:10.1006/jesp.1994.1018

Mann, N. H., & Kawakami, K. (2012). The long, steep path to equality: Progressing on egalitarian goals. *Journal of Experimental Psychology: General, 141,* 187–197. doi:10.1037/a0025602

McConahay, J. G. (1986). Modem racism, ambivalence, and the modem racism scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). New York, NY: Academic Press.

McFarland, C., Cheam, A., & Buehler, R. (2007). The perseverance effect in the debriefing paradigm: Replication and extension. *Journal of Experimental Social Psychology, 43,* 233–240. doi:10.1016/j.jesp.2006.01.010

Ministry of Justice. (2010). *Statistics on race and the criminal justice system 2010: A Ministry of Justice publication under Section 95 of the Criminal Justice Act 1991*.Retrieved from: http://www.justice.gov.uk/downloads/statistics/mojstats/stats-race-cjs-2010.pdf

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81,* 33–43. doi:10.1037//0022-3514.8I.I.33

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal of Personality and Social Psychology, 65,* 469–485. doi:10.1037//0022-3514.65.3.469

Monteith, M. J. (1996). Affective reactions to prejudice-related discrepant responses the impact of standard salience. *Personality and Social Psychology Bulletin, 22,* 48–59. doi:10.1177/0146167296221005

Monteith, M. J., Ashburn-Nardo, L., Voils, C. I., & Czopp, A. M. (2002). Putting the brakes on prejudice: On the development of cues for control. *Journal of Personality and Social Psychology, 83,* 1029–1050. doi:10.1037//0022-3514.83.5.1029

Monteith, M. J., Deneen, N. E., & Tooman, G. D. (1996). The effect of social norm activation on the expression of opinions concerning gay men and blacks. *Basic and Applied Social Psychology, 18,* 267–288. doi:10.1207/s15324834basp1803_2

Monteith, M. J., Devine, P. G., & Zuwerink, J. R. (1993). Self-directed versus other-directed affect as a consequence of prejudice-related discrepancies. *Journal of Personality and Social Psychology, 64,* 198–210. doi:10.1037//0022-3514.64.2.198

Monteith, M. J., Mark, A. Y., & Ashburn-Nardo, L. (2010). The self-regulation of prejudice: Toward understanding its lived character. *Group Processes and Intergroup Relations, 13,* 183–200. doi:10.1177/1368430209353633

Monteith, M. J., Sherman, J. W., & Devine, P. G. (1998). Suppression as a stereotype control strategy. *Personality and Social Psychology Review, 2,* 63–82. doi:10.1207/s15327957pspr0201_4

Monteith, M. J., Spicer, C. V., & Tooman, G. D. (1998). Consequences of stereotype suppression: Stereotypes on and not on the rebound. *Journal of Experimental Social Psychology, 34,* 355–377. doi:10.1006/jesp.1998.1355

Monteith, M. J., & Voils, C. I. (1998). Proneness to prejudiced responses: Toward understanding the authenticity of self-reported discrepancies. *Journal of Personality and Social Psychology, 75,* 901–916. doi:10.1037//0022-3514.75.4.901

Monteith, M. J., Voils, C. I., & Asburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial biases. *Social Cognition,* 19, 395–417. doi:10.1521/soco.19.4.395.20759

Moskowitz, G. B. (2002). Preconscious effects of temporary goals on attention. *Journal of Experimental Social Psychology, 38,* 387–404. doi:10.1016/S0022-1031(02)00001-X

Moskowitz, G. B. (2010). On the control over stereotype activation and stereotype inhibition. *Social and Personality Psychology Compass, 4,* 140–158. doi:10.1111/j.1751-9004.2009.00251.x

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

Moskowitz, G. B., & Ignarri, C. (2009). Implicit volition and stereotype control. *European Review of Social Psychology, 20,* 97–145. doi:10.1080/10463280902761896

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

Moskowitz, G. B., Li, P., Ignarri, C., & Stone, J. (2011). Compensatory cognition associated with egalitarian goals. *Journal of Experimental Social Psychology, 47,* 365–370. doi:10.1016/j.jesp.2010.08.010

Moskowitz, G. B., Salomon, A. R., & Taylor, C. M. (2000). Preconsciously controlling stereotyping: Implicitly activated egalitarian goals prevent the activation of stereotyping. *Social Cognition, 18,* 151–177. doi:10.1521/soco.2000.18.2.151

Moskowitz, G. B., & Stone, J. (2012). The proactive control of stereotype activation: Implicit goals to not stereotype. *Journal of Psychology, 22,* 172–179. doi:10.1027/2151-2604/a000110

Office of National Statistics. (2011). *Ethnicity and National Identity in England and Wales.* Retrieved from: http://www.ons.gov.uk/ons/rel/census/2011-census/key-statistics-for-local-authorities-in-england-and-wales/rpt-ethnicity.html

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45,* 867–872. doi:10.1016/j.jesp.2009.03.009

Pendry, L. F., & Macrae, N. M. (1994). Stereotypes and mental life: The case of the motivated by thwarted tactician. *Journal of Experimental Social Psychology, 30,* 303–325. doi:10.1006/jesp.1994.1015

Peruche, B. M., & Plant, E. A. (2006). Racial bias in perceptions of athleticism: The role of motivation in the elimination of bias. *Social Cognition, 24,* 438–452. doi:10.1521/soco.2006.24.4.438

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75,* 811–832. doi:10.1037//0022-3514.75.3.811

Plant, E. A., & Devine, P. G. (2001). Responses to other-imposed pro-black pressure: Acceptance or backlash? *Journal of Experimental Social Psychology, 37,* 486–501. doi:10.1006/jesp.2001.1478

Plant, E. A., & Devine, P. G. (2009). The active control of prejudice. Unpacking the intentions guiding control efforts. *Journal of Personality and Social Psychology, 96,* 640–652. doi:10.1037/a0012960

Plant, E. A., Devine, P. G., & Brazy, P. C. (2003). The bogus pipeline and motivations to respond without prejudice: Revisiting the fading and faking of racial prejudice. *Group Processes & Intergroup Relations, 6,* 187–200. doi:10.1177/1368430203006002004

Plant, E. A., Peruche, B. M., & Butz, D, A. (2005). Eliminating automatic racial bias: Making race non-diagnostic for responses to criminal suspects. *Journal of Experimental Social Psychology, 41,* 141–156. doi:10.1016/j.jesp.2004.07.004

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81,* 181–192. doi: 10.I037//0O22-3514.81.2.181

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

Quinn, K. A., & Macrae, C. N. (2005). Categorising others: The dynamics of person construal. *Journal of Personality and Social Psychology, 88,* 467–479. doi:10.1037/0022-3514.88.3.467

Quinn, K. A., Macrae, C. N., & Bodenhausen, G. V. (2003). Stereotyping and impression formation: How categorical thinking shapes person perception. In M. A. Hogg & J. Cooper (Eds.), *The SAGE handbook of social psychology* (pp. 87–109). Thousand Oaks, CA: SAGE.

Ratcliff, J. J., Lassiter, G. D., Markman, K. D., & Snyder, C. J. (2006). Gender differences in attitudes toward gay men and lesbians. The role of motivation to respond without prejudice. *Personality and Social Psychology Bulletin, 32,* 1325–1338. doi:10.1177/0146167206290213

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 38,* 119–125. doi:10.1002/1097-4679(198201)38:1<119::AID-JCLP2270380118>3.0.CO;2-I

Richeson, J.A.,& Nussbaum, R. J. (2004). The impact of multiculturalism versus color-blindness on racial bias. *Journal of Experimental Social Psychology, 40,* 417–423. doi:10.1016/j.jesp.2003.09.002

Rokeach, M. (1973). *The nature of human values.* New York: Free Press.

Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology, 5,* 856–868. doi:10.1037//0022-3514.81.5.856

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55,* 68–78. doi:10.1037//0003-066X.55.1.68

Sacchi, S., Castano, E., & Brauer, M. (2009). Perceiving one's nation: Entitativity, agency and security in the international arena. *International Journal of Psychology, 44,* 321–312. doi:10.1080/00207590802236233

Sassenberg, K., Matschke, C., Scholl, A. (2011). The impact of discrepancies from ingroup norms on group members' well-being and motivation. *European Journal of Social Psychology, 41,* 886–897. doi:10.1002/ejsp.833

Sassenberg, K. & Moskowitz, G. B. (2005). Don't stereotype, think different! Overcoming automatic stereotype activation by mindset priming. *Journal of Experimental Social Psychology, 41,* 506–514. doi:10.1016/j.jesp.2004.10.002

Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: How accessible alternatives affect goal pursuit. *Journal of Experimental Social Psychology, 38,* 368–383. doi:10.1016/S0022-1031(02)00005-7

Shah, J. Y., Friedman, R., & Kruglanski, A. W. (2002). Forgetting all else: On the antecedents and consequences of goal shielding. *Journal of Personality and Social Psychology, 83,* 1261–1280. doi:10.1037//0022-3514.83.6.1261

Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology, 76,* 482–497. doi:10.1037/0022-3514.76.3.482

Sherman, S. J., & Gorkin, L. (1980). Attitude bolstering when behavior is inconsistent with central attitudes. *Journal of Experimental Social Psychology, 16,* 388–403. doi:10.1016/0022-1031(80)90030-X

Sherman, J. W., Lee, A. Y., Bessenoff, G. R., & Frost, L. A. (1998). Stereotype efficiency reconsidered: Encoding flexibility under cognitive load. *Journal of Personality and Social Psychology, 75,* 589–606. doi:10.1037//0022-3514.75.3.589

Simon, B., & Hamilton, D. L. (1994). Self-stereotyping and social context: The effects of relative in-group size and in-group status. *Journal of Personality and Social Psychology, 66,* 699–711. doi:10.1037//0022-3514.66.4.699

Soman, D., & Cheema, A. (2004). When goals are counterproductive: The effects of violation of a behavioral goal on subsequent performance. *Journal of Consumer Research, 31,* 52–62. doi:10.1086/383423

Soper, D. S. (2014). *Effect size calculator for multiple regression.* Retrieved from: http://www.danielsoper.com/statcalc3/calc.aspx?id=5

Spencer, S. J., Fein, S., Wolfe, C. T., Fong, C., & Duinn, M. A. (1998). Automatic activation of stereotypes: The role of self-image threat. *Personality and Social Psychology Bulletin, 24,* 1139–1152. doi:10.1177/01461672982411001

Stephan, W. G. (1985). Intergroup relations. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 2, pp. 599–658). New York: Random House.

Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as efficient means of thought control. *Personality and Social Psychology Bulletin, 34,* 1332–1345. doi:10.1177/0146167208321269

STOIK Imaging. (2000). Morph Man [Computer software]. Moscow, Russia: Author.

Stroebe, W., & Insko, C. A. (1989). Stereotypes, prejudice, and discrimination: Changing conceptions in theory and research. In D. Bar-Tal, C. F. Grauman, A.W. Kruglanski, &W. Stroebe (Eds.), *Stereotypes and prejudice: Changing conceptions* (pp. 3–34). New York: Springer.

Talaska, C. A., Fiske, S. T., & Chaiken, S. (2008). Legitimating racial discrimination: Emotions, not beliefs, best predict discrimination in a meta-analysis. *Social Justice Research, 21,* 263–296. doi:10.1007/s11211-008-0071-2

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues, 25,* 79–97. doi:10.1111/j.1540-4560.1969.tb00620.x

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behaviour. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp.7–24). Chicago, IL: Nelson-Hall.

Turner, R. N., & Crisp, R. J. (2010). Explaining the relationship between ingroup identification and intergroup bias following recategorization: A self-regulation theory analysis. *Group Processes and Intergroup Relations, 13*, 251–261. doi:10.1177/1368430209351702

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 and Intergroup Relations, 6,* 37–54. doi:10.1177/1368430203006001011

Wicklund, R. A., & Gollwitzer, P. M. (1982). *Symbolic self-completion.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice: Social-cognitive goals affect amygdala and stereotype activation. *Psychological Science, 16,* 56–63. doi:10.1111/j.0956-7976.2005.00780.x

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgements and evaluations. *Psychological Bulletin, 116,* 117–142. doi:10.1037//0033-2909.116.1.117

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72,* 262–274.

Wittenbrink, B., Judd, C. M., & Park, B. (2001a). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81,* 815–827. doi:10.1037//0022-3514.81.5.815

Wittenbrink, B., Judd, C. M., & Park, B. (2001b). Evaluative versus conceptual judgements in automatic stereotyping and prejudice. *Journal of Experimental Social Psychology, 37,* 244–252. doi:10.1006/jesp.2000.1456

Wyer, N. A. (2007). Motivational influences on compliance with and consequences of instructions to suppress stereotypes. *Journal of Experimental Social Psychology, 43,* 417–424. doi:10.1016/j.jesp.2006.03.001

Wyer, N. A., Sherman, J. W., & Stroessner, S. J. (2000). The roles of motivation and ability in controlling the consequences of stereotype suppression. *Personality and Social Psychology Bulletin, 26,* 13–25. doi:10.1177/0146167200261002

Zhang, Y., Fishbach, A., & Dhar, R. (2007). When thinking beats doing: The role of optimistic expectations in goal-based choice. *Journal of Consumer Research, 34,* 567–578. doi:10.1086/520071

Zuwerink, J. R., Devine, P. G., Monteith, M. J., & Cook, D. A. (1996). Prejudice towards Blacks: With and without compunction? *Basic and Applied Social Psychology, 18,* 131–150. doi:10.1207/s15324834basp1802_2

APPENDICES

Appendix A: A More Thorough Treatment of Fishbach and Colleagues' Model

Fishbach and colleagues (Koo & Fishbach, 2008; Fishbach et al., 2011; see Figure A1) emphasise the importance of goal progress framing (whether people focus on what has already been accomplished (a to-date focus) or on what still remains to be accomplished (a to-go focus) towards a goal) in determining self-regulation throughout their research. For example, Koo and Fishbach (2008) and Fishbach et al. (2011) held discrepancy size constant at 50%. They found that when commitment is certain (as for internal goals) and people ask questions about their goal progress, a to-date focus decreases goal adherence, whereas a to-go focus increases goal adherence. In contrast, when commitment is uncertain (as for external goals) and people ask questions about their goal commitment, a to-date focus increases goal adherence, whereas a to-go focus decreases goal adherence.

However, a simplified version of Fishbach and colleagues' model (Fishbach & Dhar, 2005, 2007; Fishbach et al., 2009; Koo & Fishbach, 2008; see Figure A2) was used to guide the present research. The simplified model does not include goal progress framing as a factor. This is not to say goal progress framing is not relevant to self-regulation; research by Bonezzi, Brendl, and De Angelis (2011) accounted for discrepancy size, and goal progress framing still mattered. However, there has been no examination of the effect of both large and small discrepancies and goal progress framing within a single experiment throughout Fishbach and colleagues' research. I posit that the observed difference between the to-date and to-go progress framing in the studies conducted by Koo and Fishbach (2008) and Fishbach et al. (2011) is caused by the 50% discrepancy size manipulation rather than goal progress framing. This is based on the distinct pattern that arises when comparing the explicated model (Koo & Fishbach, 2008; Fishbach et al., 2011) where discrepancy size is controlled (50%; Figure A1)

to the simplified model (Fishbach & Dhar, 2005, 2007; Fishbach et al., 2009; Koo & Fishbach, 2008; Figure A2) where discrepancy size is manipulated.

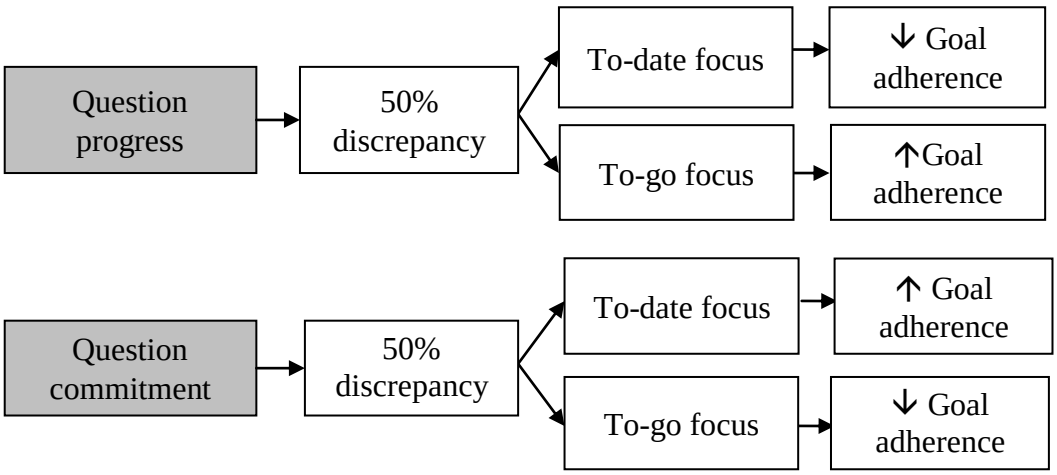

***Figure A1*. Diagrammatical representation of Fishbach and colleagues' model *including* goal progress framing.**

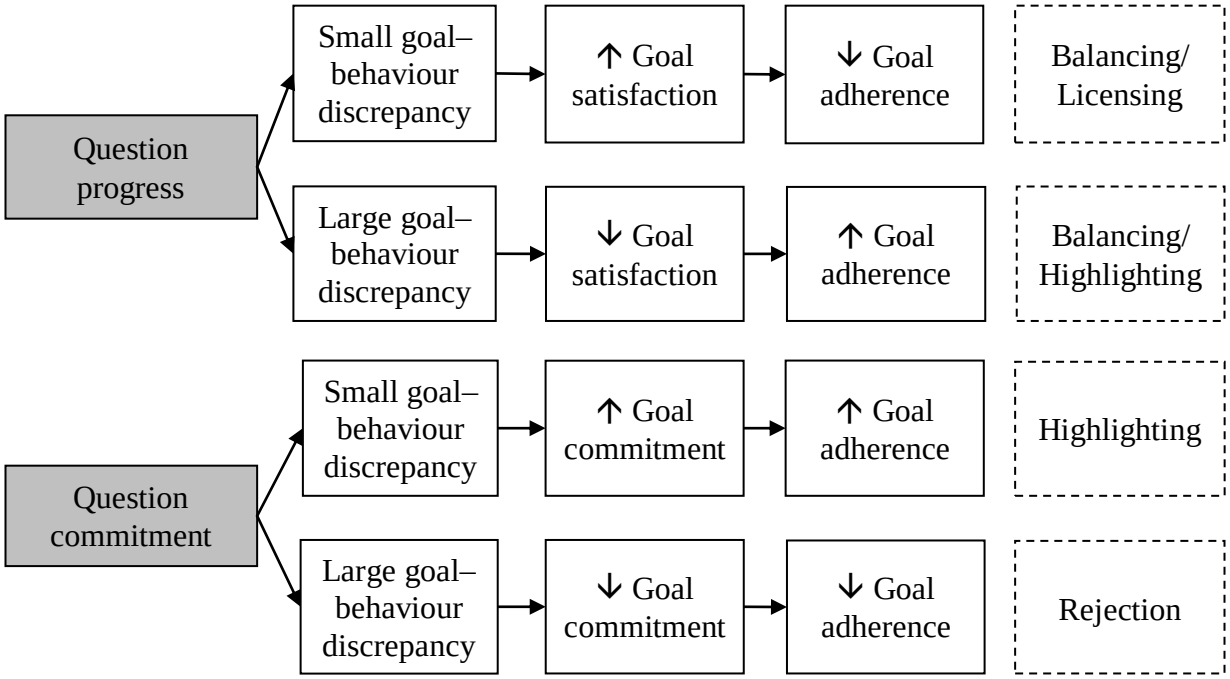

***Figure A2*. Simplified diagrammatical representation of Fishbach and colleagues' model excluding goal progress framing.**

Specifically, a 50% discrepancy coupled with a to-date frame (Figure A1), and a *small* goal–behaviour discrepancy (Figure A2), both effect goal adherence in the same manner, for both internal and external goals respectively. In contrast, a to-go frame coupled with a 50% discrepancy (Figure A1), and a *large* goal–behaviour discrepancy (A2), effect goal adherence in the same way, for both internal and external goals respectively. Thus, it appears that a to-date frame and a 50% discrepancy manipulation makes participants feel positive ("I've already done 50%"), which results in participants responding as though they received a small goal–behaviour discrepancy manipulation. For internal goals, this leads to goal relaxation, for external goals this leads to goal adherence. In contrast, a to-go frame couple with a 50% discrepancy makes participants feel negative ("I've still got 50% left"), which results in participants responding as though they received a large goal–behaviour discrepancy manipulation. For internal goals, this leads to goal adherence, for external goals this leads to goal rejection.

In addition, the effect of goal progress framing disappears when Fishbach and Dhar's (2005) studies are compared for internal goals, at least (see Table A1). Specifically, participants who focused on what they had already accomplished towards a goal (to-date frame), and who felt as though they had made progress (small goal–behaviour discrepancy) exhibited goal-inconsistent action (Fishbach & Dhar, 2005, Experiment 2). Similarly, participants who focused on what they had left to achieve a goal (to-go frame), and who felt as though they had made progress (small goal–behaviour discrepancy), also exhibited goal-inconsistent action (Fishbach & Dhar, 2005, Experiment 1).

Table A1.
A comparison of the effect of internal goals, to-date and to-go framing, and small and large goal–behaviour discrepancy size, on self-regulation across studies 1-3 for Fishbach and Dhar (2005).

Research Paper	Goal Source	Discrepancy Size	Goal Progress Framing	Outcome
Fishbach & Dhar (2005)	Internal	Small	To-date	Goal adherence decreased (Study 2 & 3)
			To-go	Goal adherence decreased (Study 1)
		Large	To-date	Goal adherence increased (Study 2)
			To-go	Goal adherence increased (Study 1)

Appendix B: Internal and External Goal Primes Used to Manipulate Goal Source in Experiment 1 (Chapter 2)

Internal Goal Focus Items

1. According to my personal values, being intolerant towards Black people is ok. (R)
2. I believe that one should be kind to all Black people.
3. I personally value being fair towards Black people.
4. It is my personal belief that Black people should NOT have an equal chance or an equal say. (R)
5. I am personally motivated to be tolerant of Black people.
6. I attempt to treat Black people as equals because I believe in equality.
7. I value being open-minded about Black people at all times.
8. I believe people should NOT be concerned about the welfare of Black people. (R)
9. I believe being biased towards Black people is ok. (R)
10. I believe that everyone should find ways to help Black people less fortunate than oneself.

External Goal Focus Items

1. People believe that others should treat Black people equally.
2. People expect others to try to hide their intolerance of Black people.
3. People value others who attempt to be open-minded about Black people.
4. People encourage others to act kindly towards Black people.
5. People expect others to be biased towards Black people. (R)
6. People believe that others should try to treat Black people unfairly. (R)
7. People value others who attempt to be tolerant of Black people.
8. People expect others to act unkindly towards Black people. (R)
9. People believe others ought to be concerned about the welfare of Black people.
10. People encourage others to treat Black people unequally. (R)

Note. (R) refers to reverse-coded items. Participants rated each statement on a 7-point scale, ranging from -3 (*disagree strongly*) to +3 (*agree strongly*).

Appendix C: Complete List of Word Stimuli Used in the LDT in Experiments 1–5 (Chapters 2 & 3)

Table C1. **Stimulus Words Used for the Practice Trials of the LDT (Experiments 1–5).**

Neutral Words	Non-Words
Calendar	Ladenrac
Compass	Mospacs
Table	Batel
Finger	Gifner

Table C2. **Stimulus Words Used for the Experimental Trials of the LDT in Blocks 1 and 2 (Experiments 1–5).**

Black-Stereotypic Words		Stereotype-Neutral Words	
Positive Words	**Negative Words**	**Positive Words**	**Negative Words**
Athletic	Aggressive	Desirable	Cockroach
Happy	Hostile	Love	Despair
Loyal	Ignorant	Paradise	Disgust
Musical	Lazy	Pleasant	Disturbing
Religious	Loud	Romance	Garbage
Rhythmic	Poor	Smile	Poison
Sensitive	Unintelligent	Vacation	Vomit

Table C3. **Stimulus Non-Words Used for the Experimental Trials of the LDT in Blocks 1 and 2 (Experiments 1–5).**

Non-Words			
Matched with Black-Stereotypic Positive Words	**Matched with Black-Stereotypic Negative Words**	**Matched with Stereotype-Neutral Positive Words**	**Matched with Stereotype-Neutral Negative Words**
Thailect	Sigresaveg	Sirabeed	Chorocack
Phapy	Shilteo	Olev	Derisap
Yolal	Gantinor	Pedarisa	Sisgudt
Calumis	Zaly	Lenapats	Ribstuding
Rigilouse	Olud	Moncare	Gerabag
Thrymich	Ropo	Lesim	Sopion
Sivenise	Tullenigentin	Tavaconi	Mitov

Appendix D: Internal and External Goal Primes Used to Manipulate Goal Source in Experiments 2– 5 (Chapter 3) and Experiment 6 (Chapter 4)

Internal Goal Focus Items

1. According to my personal values, being intolerant towards Black people is ok. (R)
2. I believe that one should be kind to all Black people.
3. I personally value being fair towards Black people.
4. It is my personal belief that Black people should NOT have an equal chance or an equal say. (R)
5. I am personally motivated to be tolerant of Black people.
6. I attempt to treat Black people as equals because I believe in equality.
7. I value being open-minded about Black people at all times.
8. I believe people should NOT be concerned about the welfare of Black people. (R)
9. I believe being biased towards Black people is ok. (R)
10. I believe that everyone should find ways to help Black people less fortunate than oneself.

External Goal Focus Items

1. Other egalitarian people believe that I should treat Black people equally.
2. Other egalitarian people would be disappointed if I were intolerant of Black people.
3. Other egalitarian people believe I should be open-minded about Black people.
4. Other egalitarian people encourage me to act kindly towards Black people.
5. Other egalitarian people would disapprove of me if I acted in a biased manner towards Black people.
6. Other egalitarian people believe that I should treat Black people fairly.
7. Other egalitarian people expect me to be tolerant of Black people.
8. Other egalitarian people would be disappointed if I treated Black people unfairly.
9. Other egalitarian people believe I ought to be concerned about the welfare of Black people.
10. Other egalitarian people would be angry if I treated Black people unequally.

Note. (R) refers to reverse-coded items. Participants rated each statement on a 7-point scale, ranging from -3 (*disagree strongly*) to +3 (*agree strongly*).

Appendix E: Complete List of Word Stimuli Used in the IAT in Experiments 2–6 (Chapters 3 & 4)

Table E1. **Stimulus Words Used during the IAT (Experiments 2–6).**

Unpleasant Words	Pleasant Words
Abuse	Diamond
Death	Happy
Disaster	Heaven
Filth	Lucky
Hatred	Miracle
Rotten	Peace
Sickness	Rainbow
Tragedy	Sunrise

Appendix F: Exploratory Measures and Analyses, Experiments 1–5

1.0 Prejudice

1.1 Rationale

The affect misattribution procedure (AMP; Payne, Cheng, Govorum, & Stewart, 2005) was included as a measure of prejudice for exploratory purposes. I wanted to explore whether the same effects would be found on the LDT and the AMP. However, as the LDT is a fairly long task (lasting approximately 4-5 minutes) the priming effect may not last long enough to influence participants' responses on the AMP.

1.2 Method

Participants completed the affect misattribution procedure (AMP) to assess their prejudice towards Black people. Participants learned that two images would flash one after another on the screen; the first image being a real-life image and the second being a Chinese pictograph. Participants were instructed to ignore the first image as this served simply as a warning signal for the Chinese pictograph. Participants were instructed to indicate whether each Chinese pictograph was more unpleasant than average using the "U" key for unpleasant, or more pleasant than average using the "P" key for pleasant. Participants completed 36 trials presented twice in a random order (a total of 72 trials). Each 36 trials included 12 trials of each type of prime (12 White faces, 12 Black faces, and 12 neutral primes (a grey square)) and 36 different Chinese pictographs. On each trial, a prime was presented for 75 ms, followed by a blank screen for 125 ms, followed by a Chinese pictograph for 100 ms. Finally a pattern mask appeared until participants responded (see Figure F1).

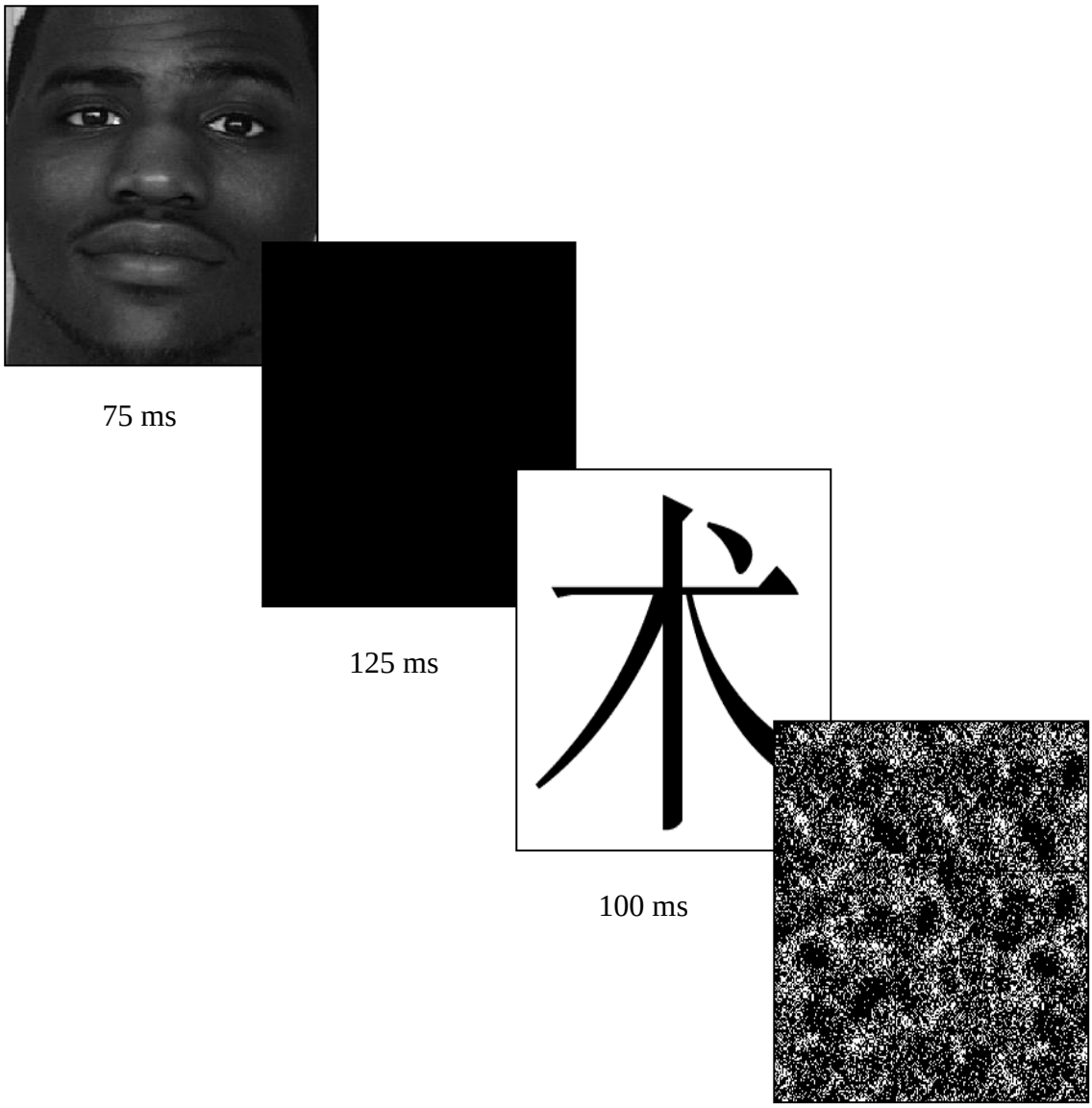

Figure F1. **AMP trial structure.**

1.3 Results

1.3.1 Experiment 1. Due to a computer malfunction, the AMP data for one participant was lost. An additional participant who could read Chinese was removed from data analysis of the AMP. The proportion of "pleasant" responses made during the AMP for each of the remaining participants was submitted to a 2 (Goal Source: internal vs. external) × 2 (Prime

Race: White vs. Black) mixed-model ANOVA with goal source as a between-participants factor. No main effects or interactions were significant, all $p > .16$; goal source had no effect.

1.3.2 Experiment 2. Due to a computer malfunction, the AMP data for one participant was lost. The proportion of "pleasant" responses made during the AMP for each of the remaining participants was submitted to a 2 (Goal Source: internal vs. external) × 2 (Prime Race: White vs. Black) mixed-model ANOVA with goal source as a between-participants factor. No main effects or interactions were significant, all $p > .09$; goal source had no effect.

1.3.3 Experiment 3. Two participants were removed from data analysis, including one participant who could read Chinese and one participant who satisficed (i.e., responded "pleasant" to all of the Chinese pictographs). The proportion of "pleasant" responses made during the AMP for each of the remaining participants was submitted to a 2 (Goal Source: internal vs. external) × 2 (Prime Race: White vs. Black) mixed-model ANOVA with goal source as a between-participants factor. The analysis revealed a significant main effect of prime race, $F(1, 79) = 4.15$, $p = .045$, $\eta^2_p = .05$, indicating that participants evaluated the Chinese pictographs as more pleasant following a White prime ($M = 0.64$, $SE = 0.02$) compared to a Black prime ($M = 0.59$, $SE = 0.02$). No additional main effects or interactions were significant, all $p > .56$; goal source had no effect.

1.3.4 Experiment 4. Five participants were removed from data analysis, including one participant who could read Chinese and four participants who satisficed (i.e., responded "pleasant" to all of the Chinese pictographs). The proportion of "pleasant" responses made during the AMP for each of the remaining participants was submitted to a 2 (Goal Source: internal vs. external) × 2 (Prime Race: White vs. Black) mixed-model ANOVA with goal source as a between-participants factor. No main effects or interactions were significant, all $p > .10$; goal source had no effect.

1.3.5 Experiment 5. Due to a computer malfunction, the AMP data for two participants were lost. A further two participants were removed from data analysis for satisficing (i.e., responded "pleasant" to all of the Chinese pictographs). The proportion of "pleasant" responses made during the AMP for each of the remaining participants was submitted to a 2 (Goal Source: internal vs. external) × 2 (Goal–Behaviour Status: goal–behaviour discrepancy vs. goal–behaviour match) × 2 (Prime Race: White vs. Black) mixed-model ANOVA with goal source and goal–behaviour status as a between-participants factors. The analysis revealed a significant main effect of goal source, $F(1, 94) = 6.83$, $p = .010$, $\eta^2_p = .07$, indicating that the overall proportion of "pleasant" responses was higher in the internal goal condition ($M = 0.67$ $SE = 0.02$) than the external goal condition ($M = 0.58$, $SE = 0.02$). No additional main effects or interactions were significant, all $p > .066$.

2.0 Chronic Motivation

2.1 Rationale

Both temporary egalitarian goals and chronic egalitarian motivations (both internal and external) have been examined within the stereotyping and prejudice literature. Research has demonstrated that both temporarily primed internal goals and chronic internal motivation result in prejudice-control (e.g., Fehr & Sassenberg, 2010; Mann & Kawakami, 2011), and temporarily primed internal goals result in stereotype-control (e.g., Moskowitz & Li, 2011; Moskowitz & Stone, 2012) after a self-regulatory failure. Research examining external motivation is limited, however. Only one investigation by Fehr and Sassenberg (2010) has examined how self-regulatory failure effects prejudice for chronically externally motivated participants. They found no effect of chronic external motivation on prejudice following a self-regulatory failure compared to no self-regulatory failure.

Little research has really examined the potential interplay between temporary and chronic egalitarian motivations. One recent study by Legault et al. (2011) has examined the impact of temporarily priming internal and external motivation to act nonprejudiced on prejudice activation (measured via the IAT; Greenwald et al., 1998) and prejudice expression (measured via the symbolic racism scale; Henry & Sears, 2002). They also examined the potential *mediating* role of chronic autonomous motivation to act nonprejudiced (measured via their *Motivation to be Nonprejudiced* scale). They found that the temporarily primed internal goal condition resulted in less prejudice activation and expression, whereas the temporarily primed external goal condition resulted in greater prejudice activation and expression. However, this effect was partially mediated by autonomous motivation, such that the effect of the temporarily primed internal goal condition on prejudice activation and expression was boosted by autonomous motivation, whereas the effect of temporarily primed external condition seemed to be relatively independent of autonomous motivation. This finding might imply that a match between the source of a temporary goal and a chronic goal an additive effect, at least for internal goals. Indeed, there is some evidence for the importance of self-regulatory fit in Higgins' (1997, 1998) regulatory focus theory: that promotion people are more motivated in promotion-framed than prevention-framed contexts, whereas prevention people show the opposite.

I was interested in examining the potential *moderating* role of chronic goals to act egalitarian on the impact of the temporarily primed egalitarian goals on stereotype activation. Specifically, I was interested in what would happen when there is a mismatch between the source of a chronic goal and temporarily primed goal. On the one hand, participants may respond consistently with their chronic motivations rather than the temporarily primed goals (i.e., chronic goals takes priority over temporarily primed goals). Alternatively, participants

may respond consistently with the temporarily primed goals rather than their chronic motivations (i.e., temporarily primed goals take priority over temporary goals). Depending on which motivation is most salient at that moment in time may determine which motivation effects behaviour.

I included Plant and Devine's (1998) internal and external motivation to act nonprejudiced scales as my own temporary internal and external egalitarian goal primes were adapted from these scales. I also included Sheldon and Elliot's (1999) measure of self-concordance for two reasons. First, there are clear parallels between Plant and Devine's (1998) internal and external motivation to act nonprejudiced scales. The self-concordance model proposes that self-concordant goals are pursued for intrinsic reasons (i.e., because of the fun and enjoyment pursuing the goal provides) or identified reasons (i.e., because the goal is important). Both internal goals and self-concordant goals are freely chosen, reflect a person's values and beliefs, and are self-defining (Plant & Devine, 1998; Sheldon & Elliot, 1999). In contrast, goals that are not self-concordant are pursued for introjected (i.e., to avoid feeling guilty, ashamed, or anxious) or extrinsic (i.e., because other people/the situation demands it, or to achieve rewards (e.g., monetary rewards)) reasons. Both external goals and non-self-concordant goals are externally imposed, do not reflect a person's values and beliefs well, and are not self-defining (Plant & Devine, 1998; Sheldon & Elliot, 1999). Second, Plant and Devine's (1998) chronic motivation scales assess internal and external motivation to act *nonprejudiced*, whereas my temporary goal primes induced either an internal or external *egalitarian* goal. Sheldon and Elliot's (1999) self-concordance measure was easily adaptable to measure chronic egalitarian motivation.

2.2 Chronic Internal and External Motivation to Act Nonprejudiced

2.2.2 Method. Participants completed both Plant and Devine's (1998) 5-item IMS scale measuring internal motivation to respond nonprejudiced (e.g., "I attempt to act in nonprejudiced ways towards Black people because it is personally important to me") and 5-item EMS scale measuring external motivation to respond nonprejudiced (e.g., "Because of today's PC (politically correct) standards I try to appear nonprejudiced towards Black people"). Participants rated how much they agreed or disagreed with each statement, using a 9-point scale ranging from 0 (*disagree strongly*) to 8 (*agree strongly*). The items from both scales were intermixed and presented in a randomised order[41].

2.2.3 Results.

2.2.3.1 Experiment 2. After reverse-coding appropriate items, indices of IMS (M = 5.99, SD = 1.32, α = .86) and EMS (M = 3.60, SD = 1.83, α = .88) were created by averaging responses, with higher scores indicating greater internal or external motivation, respectively (see Plant & Devine, 1998). Interestingly, the IMS and EMS scales were negatively correlated, $r(119) = -.21$, $p = .025$), inconsistent with previous research that has shown the IMS and EMS to be uncorrelated (e.g., Plant & Devine, 1998; Devine et al., 2002).

The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × IMS × EMS standardised regression analysis. The analysis revealed a significant interaction between Goal Source × IMS × EMS, $\beta = 29.96$, $t(111) = 2.84$, $p = .005$, $R^2 = .12$ (interaction means are presented in Figure F2). A post-hoc power analysis (G*Power; Faul et al., 2009) using effect size F^2 (0.13; determined using Soper's (2014) effect size conversion calculator), critical alpha (.05), total sample size (119

[41] The IMS and EMS measure followed the identification with egalitarians measure in Experiments 2–5. Participants completed the IMS and EMS along with the contact, social desirability, and reactance measures (the order of presentation was randomised for these four measures).

participants), and number of predictors (3), indicated that the achieved statistical power (1 - β) was 0.92, which exceeds the recommended .80 (Cohen, 1988). As a result, the interaction was decomposed by examining IMS and EMS one standard deviation above and below the mean.

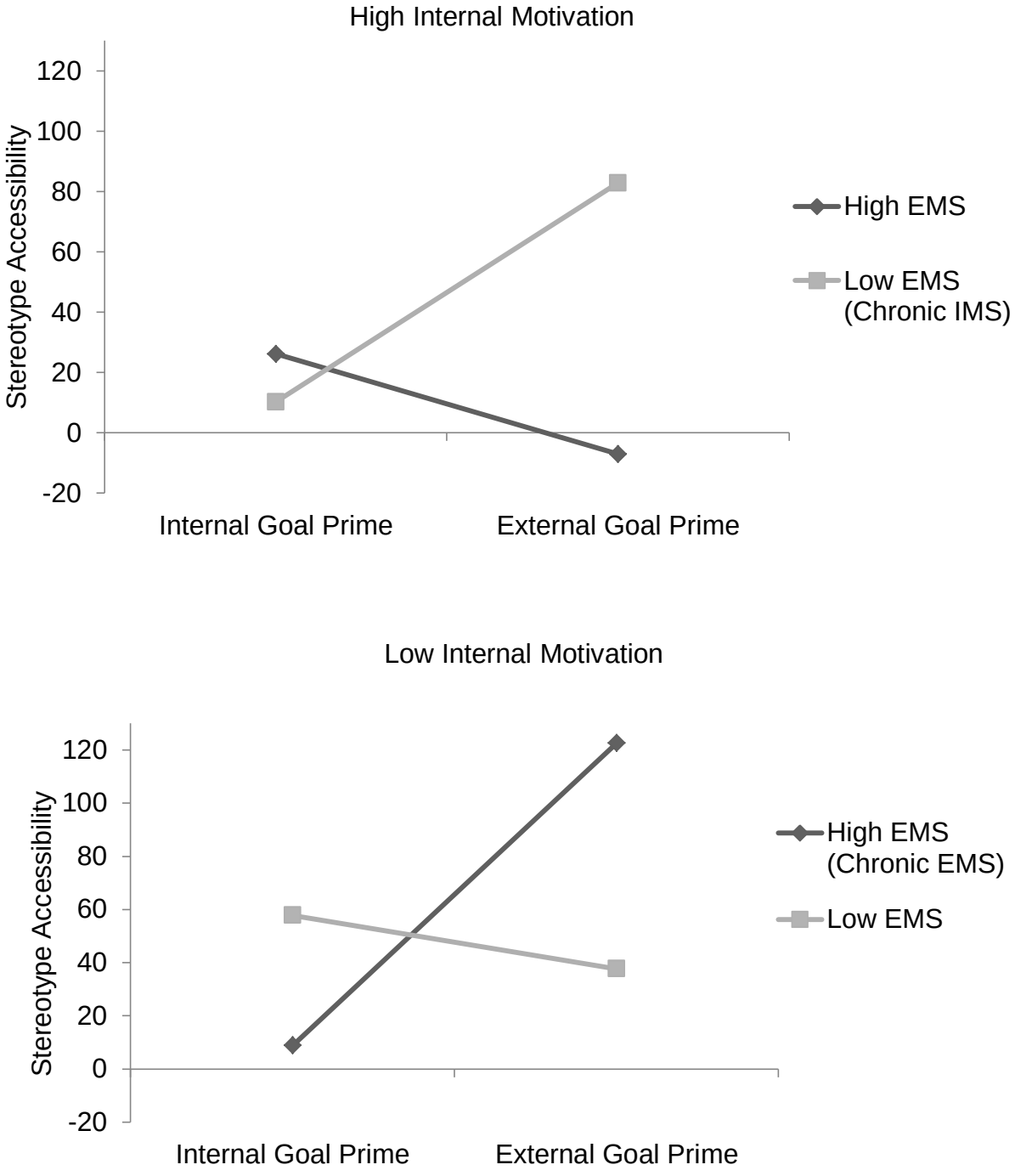

Figure F2. **Stereotype accessibility (higher scores indicate greater accessibility of Black stereotypes) as a function of chronic internal and external motivation to act nonprejudiced, and temporary goal source, Experiment 2.**

2.2.3.1.1 High IMS. For participants high in both chronic internal and external motivation, the goal source priming manipulation had no effect on stereotype accessibility, β = 16.59, $t(111) = 0.71$, $p = .48$, $R^2 = .12$. However, for participants high in chronic internal motivation (high in internal and low in external motivation) the goal source priming manipulation had a significant effect on stereotype accessibility, $\beta = -36.25$, $t(111) = 2.16$, $p = .033$, $R^2 = .12$. Specifically, these individuals exhibited greater accessibility of Black stereotypes in the external than the internal goal prime condition.

2.2.3.1.2 Low IMS. For participants low in both internal and external motivation, the goal source priming manipulation had no effect on stereotype accessibility, $\beta = 10.04$, $t(111) = 0.46$, $p = .65$, $R^2 = .12$. However, for participants high in chronic external motivation (low in internal and high in external motivation) the goal source priming manipulation had a significant effect on stereotype accessibility, $\beta = -56.95$, $t(111) = 2.78$, $p = .006$, $R^2 = .12$. Specifically, these individuals exhibited greater accessibility of Black stereotypes in the external than the internal goal prime condition.

2.2.3.2 Experiment 3. After reverse-coding appropriate items, indices of IMS (M = 6.27, SD = 1.17, $\alpha = .74$) and EMS (M = 3.31, SD = 1.84, $\alpha = .87$) were created by averaging responses, with higher scores indicating greater internal or external motivation, respectively (see Plant & Devine, 1998). As in Experiment 2, the IMS and EMS scales were negatively correlated, $r(83) = -.36$, $p = .001$), again inconsistent with previous research that has shown the IMS and EMS to be uncorrelated (e.g., Plant & Devine, 1998; Devine et al., 2002).

The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × IMS × EMS standardised regression analysis. The analysis revealed a significant main effect of chronic internal motivation, $\beta = 27.76$, $t(74) = 2.17$, $p = .034$, $R^2 = .10$, indicating that as internal motivation increased, stereotype accessibility

increased. No other main or interaction effects were significant, however, all $p > .14$. A post-hoc power analysis (G*Power; Faul et al., 2009) using effect size F^2 (0.10); determined using Soper's (2014) effect size conversion calculator), critical alpha (.05), total sample size (82 participants), and number of predictors (3), indicated that the achieved statistical power (1 - β) was 0.67, which is lower than the recommended .80 (Cohen, 1988), but still adequate.

2.2.3.3 Experiment 4. After reverse-coding appropriate items, indices of IMS (M = 6.08, SD = 1.40, α = .79) and EMS (M = 3.13, SD = 1.97, α = .90) were created by averaging responses, with higher scores indicating greater internal or external motivation, respectively (see Plant & Devine, 1998). Unlike in the previous studies reported in this thesis, the IMS and EMS scales were uncorrelated, $r(58) = -.08$, $p = .56$), consistent with previous research (e.g., Plant & Devine, 1998; Devine et al., 2002).

The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × IMS × EMS standardised regression analysis. The analysis revealed a significant main effect of chronic external motivation, $\beta = -22.92$, $t(49) = 2.07$, $p = .044$, $R^2 = .27$, indicating that as external motivation increased, stereotype accessibility decreased. The analysis also revealed two significant interactions. First, a significant Goal Source × IMS interaction, $\beta = 27.97$, $t(49) = 2.86$, $p = .006$, $R^2 = .27$. Interaction means are presented in Figure F3. The interaction was decomposed by examining IMS one standard deviation above and below the mean. For participants high in IMS, the goal source priming manipulation had no effect on stereotype accessibility, $\beta = 16.64$, $t(49) = 1.26$, $p = .22$, $R^2 = .27$. However, or participants low in IMS the goal source priming manipulation had a significant effect on stereotype accessibility, $\beta = -39.30$, $t(49) = 2.73$, $p = .009$, $R^2 = .27$. Specifically, these individuals exhibited greater accessibility of Black stereotypes in the external than the internal goal prime condition.

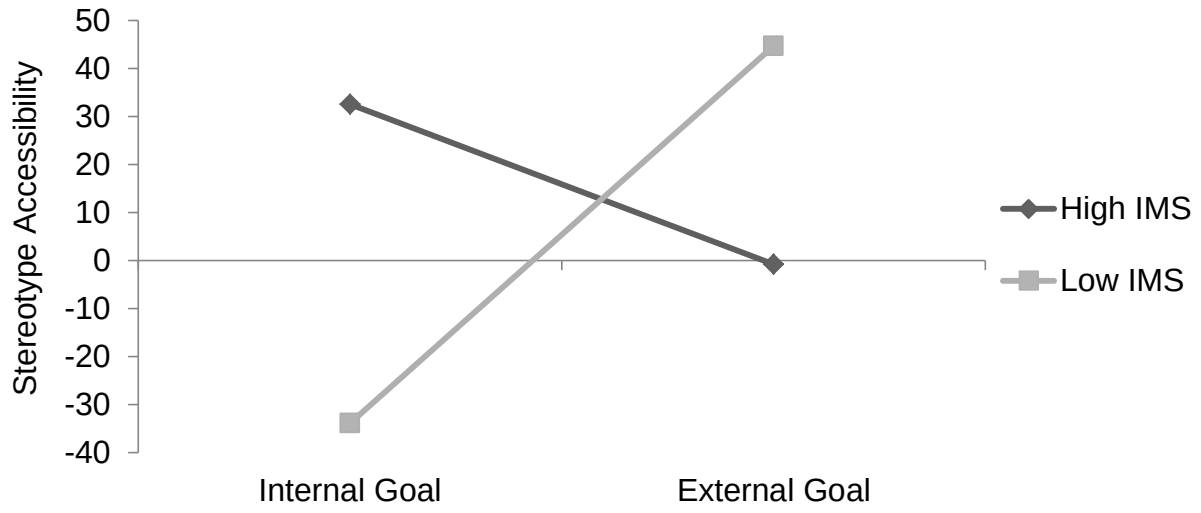

Figure F3. **Stereotype accessibility (higher scores indicate greater accessibility of Black stereotypes) as a function of IMS and goal source, Experiment 4.**

Second, a significant IMS × EMS interaction, $\beta = 32.41$, $t(49) = 3.38$, $p = .001$, $R^2 = .27$. Interaction means are presented in Figure F4. The interaction was decomposed by examining EMS one standard deviation above and below the mean. For participants high in IMS, those participants high in EMS exhibited greater stereotype accessibility than participants low in EMS, $\beta = 37.67$, $t(49) = 2.37$, $p = .022$, $R^2 = .27$. For participants low in IMS, those participants low in EMS exhibited greater stereotype accessibility than participants high in EMS, $\beta = -27.14$, $t(49) = 2.45$, $p = .018$, $R^2 = .27$.

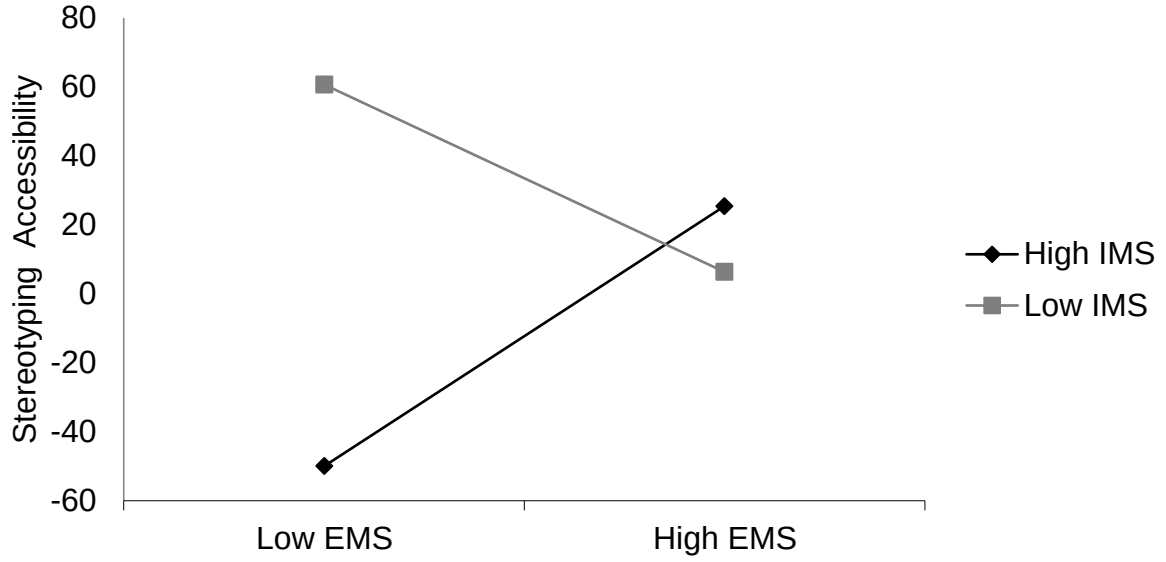

Figure F4. **Stereotype accessibility (higher scores indicate greater accessibility of Black stereotypes) as a function of IMS and EMS, Experiment 4.**

No other main effects or interactions were significant, all $p > .067$. A post-hoc power analysis (G*Power; Faul et al., 2009) using effect size F^2 (0.38); determined from the R^2 using Soper's (2014) effect size conversion calculator), critical alpha (.05), total sample size (57 participants), and number of predictors (3), indicated that the achieved statistical power (1 - β) was 0.97, which exceeds the recommended .80 (Cohen, 1988).

2.2.3.4 Experiment 5. After reverse-coding appropriate items, indices of IMS (M = 6.51, SD = 1.25, α = .82) and EMS (M = 3.48, SD = 1.95, α = .89) were created by averaging responses, with higher scores indicating greater internal or external motivation, respectively (see Plant & Devine, 1998). As in Experiments 2 and 3, the IMS and EMS scales were negatively correlated, $r(102) = -.34$, $p = .001$, inconsistent with previous research that has shown the IMS and EMS to be uncorrelated (e.g., Plant & Devine, 1998; Devine et al., 2002).

The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Goal–Behaviour Status (effects coded as -1 goal–behaviour discrepancy and +1 goal–behaviour match) × IMS × EMS standardised regression analysis.

No main or interaction effects were significant, all $p > .10$ A post-hoc power analysis (G*Power; Faul et al., 2009) using effect size F^2 (0.19); determined from the R^2 using Soper's (2014) effect size conversion calculator), critical alpha (.05), total sample size (99 participants), and number of predictors (3), indicated that the achieved statistical power (1 - β) was 0.96, which exceeds the recommended .80 (Cohen, 1988).

2.2.3 Summary.

In the present thesis, chronic internal and external motivation to act nonprejudiced only had a moderating role on the impact of the temporarily primed egalitarian goals on stereotype activation in Experiment 2. Experiment 2 demonstrated that chronically internally motivated (high in internal and low in external motivation) and chronically externally motivated (low in internal and high in external motivation) participants *both* exhibited accessibility of Black stereotypes after being primed with an external egalitarian goal, but little to no stereotype accessibility after being primed with an internal egalitarian goal. This finding suggests that participants may have prioritised the temporarily primed goals above their own chronic motivations, potentially because the temporarily primed goals were more salient within the experimental context.

When the temporarily primed goal and participants' chronic motivation matched, participants responded consistently with Fishbach and colleagues' model. Chronically internally motivated participants exhibited little to no stereotype accessibility in the temporarily primed internal goal condition, consistent with Fishbach and colleagues' model for internal goals. Whereas, chronically externally motivated participants exhibited stereotype accessibility in the temporarily primed external goal condition, consistent with Fishbach and colleagues' model for external goals.

When the temporarily primed goal and participants' chronic motivation mismatched, participants responded consistently with Fishbach and colleagues' model, but only in relation to the temporarily primed goals. This suggests that participants prioritised the temporarily primed goal above their own chronic motivation. Chronically internally motivated participants exhibited stereotype accessibility in the temporarily primed external goal condition, consistent with Fishbach ad colleagues' model for external goals. Whereas, chronically externally motivated participants exhibited little to no stereotype accessibility in the temporarily primed internal goal condition, consistent with Fishbach and colleagues' model for internal goals.

2.3 Self-Concordance

2.3.1 Method.

2.3.1.1 Experiments 2 & 3. Participants completed Sheldon and Elliot's (1999) measure of self-concordance to assess participants' motivation for goal pursuit. Participants read four statements, each describing a different reason for acting egalitarian based on self-concordance theory (Sheldon & Elliot, 1999). The external reason read, "you are egalitarian because somebody else wants you to or because the situation demands it". The introjected reason read, "you are egalitarian because you would feel ashamed, guilty or anxious if you didn't". The identified reason read, "you are egalitarian because you really believe it's an important goal to have". Finally, the intrinsic reason read, "you are egalitarian because of the fun and enjoyment that it provides you". Participants indicated the extent to which they act egalitarian because of each reason, using a 9-point scale ranging from 0 (*not at all for this reason*) to 8 (*completely for this reason*). The items were presented in a randomised order[42].

2.3.1.2 Experiments 4 & 5. The measure of self-concordance was identical to the self-concordance measure used in Experiment 2 & 3, except some minor alterations to the wording

[42] The self-concordance measure was presented along with the measures of goal commitment and goal importance (the order of presentation was randomised for these three measures).

of the reasons were made so that "I" was used instead of "you" to make responding to the items easier.

2.3.2 Results.

2.3.2.1 Experiment 2. An index of self-concordance was created by subtracting the sum of participants' responses to the introjected and external scores from the sum of participants' responses to identified and intrinsic scores, with higher scores indicating greater autonomous motivation (M = 2.54, SD = 4.62; see Sheldon & Elliot, 1999). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Self-Concordance standardised regression analysis. The analysis revealed a significant main effect of goal source, β = -22.80, $t(115) = 2.36$, $p = .020$, $R^2 = .05$; accessibility of Black stereotypes was greater in the external than the internal goal condition. No additional main effects or interactions were significant, all $p > .48$; self-concordance had no effect.

2.3.2.2 Experiment 3. An index of self-concordance was created by subtracting the sum of participants' responses to the introjected and external scores from the sum of participants' responses to identified and intrinsic scores, with higher scores indicating greater autonomous motivation (M = 2.54, SD = 4.44; see Sheldon & Elliot, 1999). The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Self-Concordance standardised regression analysis. No main effects or interactions were significant, all $p > .06$; self-concordance had no effect.

2.3.2.3 Experiment 4. An index of self-concordance was created by subtracting the sum of participants' responses to the introjected and external scores from the sum of participants' responses to identified and intrinsic scores, with higher scores indicating greater autonomous motivation (M = 3.22, SD = 3.91; see Sheldon & Elliot, 1999). The negative

stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Self-Concordance standardised regression analysis. No main effects or interactions were significant, all $p > .56$; self-concordance had no effect.

2.3.2.4 Experiment 5. An index of self-concordance was created by subtracting the sum of participants' responses to the introjected and external scores from the sum of participants' responses to identified and intrinsic scores, with higher scores indicating greater autonomous motivation ($M = 4.28$, $SD = 4.32$; see Sheldon & Elliot, 1999). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Goal–Behaviour Status (effects coded as -1 goal–behaviour discrepancy and +1 goal–behaviour match) × Self-Concordance standardised regression analysis. No main effects or interactions were significant, all $p > .06$; self-concordance had no effect.

3.0 Contact

3.1 Rationale

Voci and Hewstone's (2003) measure of the quantity and quality of contact was included for exploratory purposes. To the best of my knowledge, no research has examined whether contact moderates the influence of goal source on stereotype accessibility. Yet, past research has argued that the likelihood of stereotype activation and use decreases as the amount of personalised contact with a target increases (Brewer, 1996). I wanted to explore whether participants who have more contact (more frequent and higher quality) might demonstrate less stereotype accessibility regardless of the goal source and goal–behaviour discrepancy manipulations. In contrast, participants who have less contact (less frequent and lower quality) might be influenced by the goal source manipulation.

3.2 Method

To assess how much contact participants had with Black people, participants completed Voci and Hewstone's (2003) measure of the quantity and quality of contact. To assess the quantity of contact, participants indicated how many Black people (and separately, Black students) they know, on a 5-point scale ranging from *Zero* to *More than 10*, and how frequently they have contact with Black people (and separately, Black students), on a 5-point scale ranging from 0 (*never*) to 4 (*extremely frequently*).To assess the quality of contact, participants indicated whether their contact with Black students was pleasant, cooperative, and superficial, on a 5-point scale ranging from 0 (*not at all*) to 4 (*extremely*).

3.3 Results

3.3.1 Experiment 2. After reverse coding of appropriate items, two indices were created. First, an index of the quantity of contact with Black people and Black students was created by averaging responses, with higher scores indicating higher quantity of contact (M = 2.08, SD = 0.99, α = .87). Second, an index of the quality of contact with Black students was created by averaging responses, with higher scores indicating higher quality of contact (M = 3.11, SD = 0.58, α = .60). As per Voci and Hewstone (2003), the two indices assessing the quantity of contact and quality of contact were multiplied together to create a single index of contact (M = 6.66, SD = 3.77).

The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Contact standardised regression analysis. The analysis revealed a significant main effect of goal source, β = -22.37, $t(115) = 2.32$, $p = .022$, $R^2 = .06$; accessibility of Black stereotypes was greater in the external than the internal goal condition. No additional main effects or interactions were significant, all $p > .44$; contact had no effect.

3.3.2 Experiment 3. An index of the quantity of contact with Black people and Black students was created by averaging responses, with higher scores indicating higher quantity of contact (M = 2.23, SD = .82, α = .77). After reverse coding appropriate items, the three items assessing the quality of contact were not indexed due to low internal consistency (α = .27). Consequently, the negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Contact Quantity standardised regression analysis. No main effects or interactions were significant, all p > .28; contact quantity had no effect.

Additionally, each item of the quality of contact measure was assessed individually in order to examine the quality of contact. The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Contact Pleasantness standardised regression analyses, which revealed no significant effects or interactions, all p > .27. The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Contact Cooperativeness standardised regression analyses, which revealed no significant effects or interactions, all p > .65. The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Contact Superficiality standardised regression analyses, which revealed no significant effects or interactions, all p > .71.

3.3.3 Experiment 4. An index of the quantity of contact with Black people was created by averaging responses, with higher scores indicating higher quantity of contact (M = 2.32, SD = .90, α = .87). After reverse coding appropriate items, the three items assessing the quality of contact were not indexed due to poor internal consistency (α = .48). Consequently, the negative stereotype accessibility index was submitted to a Goal Source (effects coded as -

1 for external and +1 for internal) × Contact Quantity standardised regression analysis. No main effects or interactions were significant, all $p > .58$; contact quantity had no effect.

Additionally, each item of the quality of contact measure was assessed individually in order to examine the quality of contact. The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Contact Pleasantness standardised regression analyses. There was only a significant Goal Source × Contact Pleasantness interaction, $\beta = 21.70$, $t(53) = 2.22$, $p = .031$, $R^2 = .11$. Interaction means are presented in Figure F5. However, when the interaction was decomposed by examining Pleasant ratings one standard deviation above and below the mean, neither simple slope was significant. For participants who rated their contact as highly pleasant, the goal source priming manipulation had no effect on stereotype accessibility, $\beta = 18.61$, $t(53) = 1.34$, $p = .19$, $R^2 = .11$. For participants who rated their contact as lowly pleasant, the goal source priming manipulation had no effect on stereotype accessibility, $\beta = -24.79$, $t(53) = 1.84$, $p = .071$, $R^2 = .11$. No additional main effects or interactions were significant, all $p > .34$.

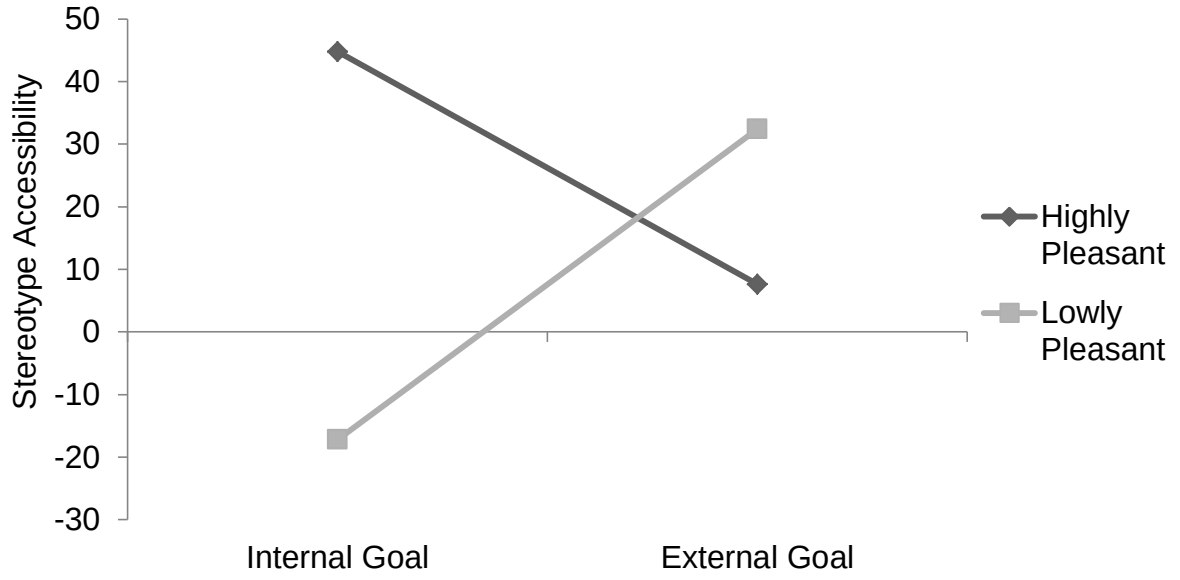

Figure F5. **Stereotype accessibility (higher scores indicate greater accessibility of Black stereotypes) as a function of contact pleasantness and goal source, Experiment 4.**

Furthermore, the negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Contact Cooperativeness standardised regression analyses, which revealed no significant effects or interactions, all $p > .12$. The negative stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Contact Superficiality standardised regression analyses, which revealed no significant effects or interactions, all $p > .15$.

3.3.4 Experiment 5. After reverse coding of appropriate items, two indices were created. First, an index of the quantity of contact with Black people and Black students was created by averaging responses, with higher scores indicating higher quantity of contact (M = 2.06, SD = .93, α = .82). Second, an index of the quality of contact with Black students was created by averaging responses, with higher scores indicating higher quality of contact (M = 3.05, SD = .60, α = .59). As per Voci and Hewstone (2003), the two indices assessing the quantity of contact and the quality of contact were multiplied together to create a single index

of contact (M = 6.49, SD = 3.43). The stereotype accessibility index was submitted to a Goal Source (effects coded as -1 for external and +1 for internal) × Goal–Behaviour Status (effects coded as -1 goal–behaviour discrepancy and +1 goal–behaviour match) × Contact standardised regression analysis. No main effects or interactions were significant, all $p > .11$; contact had no effect.

4.0 Social Desirability

4.1 Rationale

Reyonold's (1982) 13-item short form C of the Marlowe-Crowne social desirability scale was included to ensure that participants were not responding in a socially desirable manner, particularly participants in the external goal condition who should be concerned with hiding nonegalitarian behaviour from others. To address this concern, participants completed the short form C of the Marlowe-Crowne social desirability scale. I subsequently examined whether participants' propensity for responding in a socially desirable manner differed significantly between the internal and external goal conditions in Experiments 2–5.

4.2 Method

To assess whether participants were responding in a socially desirable manner, participants completed Reynolds' (1982) 13-item short form C of the Marlowe-Crowne social desirability scale. Participants indicated whether 13 randomly presented statements (e.g., "I'm always willing to admit it when I make a mistake") were true or false.

4.3 Results

4.3.1 Experiment 2. After reverse-coding, an index of social desirability was created by summing socially desirable responses, with higher scores indicating higher socially desirable responding (α = .56; see Crowne-Marlow, 1960; Reynolds, 1982). Overall, the level of socially desirable responding (M = 5.66, SD = 2.36) was similar to previous research

investigating the Marlowe-Crowne short form C (e.g., Loo & Loewen, 2004; Loo & Thorpe, 2000). A one-way between-participants ANOVA indicated that socially desirable responding did not differ between the internal and external goal conditions, $F(1, 117) = 0.52$, $p = .47$, $\eta^2_p < .01$.

4.3.2 Experiment 3. After reverse-coding, an index of social desirability was created by summing socially desirable responses, with higher scores indicating higher socially desirable responding ($\alpha = .65$; see Crowne-Marlow, 1960; Reynolds, 1982). Overall, the level of socially desirable responding ($M = 5.86$, $SD = 2.63$) was similar to previous research investigating the Marlowe-Crowne short form C (e.g., Loo & Loewen, 2004; Loo & Thorpe, 2000). A one-way between-participants ANOVA indicated that socially desirable responding did not differ between the internal and external goal conditions, $F(1, 81) = 0.44$, $p = .51$, $\eta^2_p < .01$.

4.3.3 Experiment 4. After reverse-coding, an index of social desirability was created by summing socially desirable responses, with higher scores indicating higher socially desirable responding ($\alpha = .57$; see Crowne-Marlow, 1960; Reynolds, 1982). Overall, the level of socially desirable responding ($M = 5.74$, $SD = 2.43$) was similar to previous research investigating the Marlowe-Crowne short form C (e.g., Loo & Loewen, 2004; Loo & Thorpe, 2000). A one-way between-participants ANOVA indicated that socially desirable responding did not differ between the internal and external goal conditions, $F(1, 56) = 1.36$, $p = .25$, $\eta^2_p = .02$.

4.3.4 Experiment 5. After reverse-coding, an index of social desirability was created by summing socially desirable responses, with higher scores indicating higher socially desirable responding ($\alpha = .70$; see Crowne-Marlow, 1960; Reynolds, 1982). Overall, the level of socially desirable responding ($M = 6.20$, $SD = 2.78$) was similar to previous research

investigating the Marlowe-Crowne short form C (e.g., Loo & Loewen, 2004; Loo & Thorpe, 2000). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for social desirability bias was conducted. No main effects or interactions were significant, all $p > .47$.

5.0 Reactance

5.1 Rationale

Dowd, Milne, and Wise's (1991) therapeutic reactance scale was included to ensure that any stereotype activation exhibited by participants in the external goal condition was not the result of a general predisposition to respond with reactance generally. To address this concern, participants completed the therapeutic reactance scale. I subsequently examined whether participants' general propensity for reactance differed significantly between the internal and external goal conditions in Experiments 2–5.

5.2 Method

To assess reactance, participants completed Dowd, Milne, and Wise's (1991) therapeutic reactance scale. Participants rated how much they agreed or disagreed with 28 randomly presented statements (e.g., "I resent authority figures who try to tell me what to do"), on a 4-point scale ranging from 1 (*disagree strongly*) to 4 (*agree strongly*).

5.3 Results

5.3.1 Experiment 2. After reverse-coding appropriate items, an index of reactance was created by summing responses, with higher scores indicating higher reactance ($\alpha = .83$; see Dowd et al., 1991). Participants in the present experiment exhibited similar reactance scores ($M = 64.24$ (out of 112 in total), $SD = 7.98$, range = 44 to 89) to the normative data provided by Dowd et al. (1991; $M = 68.86$ (out of 112 in total), $SD = 6.59$, range = 46 to 83). A one-

way between-participants ANOVA indicated that reactance did not differ between the internal and external goal conditions, $F(1, 117) = 0.22$, $p = .64$, $\eta^2_p < .01$.

5.3.2 Experiment 3. After reverse-coding appropriate items, an index of reactance was created by summing responses, with higher scores indicating higher reactance ($\alpha = .80$; see Dowd et al., 1991). Participants exhibited similar reactant scores ($M = 64.81$ (out of 112 in total), $SD = 7.56$, range = 43 to 84) to the normative data provided by Dowd et al. (1991; $M = 68.86$ (out of 112 in total), $SD = 6.59$, range = 46 to 83). A one-way between-participants ANOVA indicated that reactance did not differ between the internal and external goal conditions, $F(1, 81) = 2.06$, $p = .16$, $\eta^2_p = .03$.

5.3.3 Experiment 4. After reverse-coding appropriate items, an index of reactance was created by summing responses, with higher scores indicating higher reactance ($\alpha = .83$; see Dowd et al., 1991). Participants exhibited similar reactant scores ($M = 67.03$ (out of 112 in total), $SD = 8.29$, range = 49 to 92) to the normative data provided by Dowd et al. (1991; $M = 68.86$ (out of 112 in total), $SD = 6.59$, range = 46 to 83). A one-way between-participants ANOVA indicated that reactance did not differ between the internal and external goal conditions, $F(1, 56) = 0.12$, $p = .73$, $\eta^2_p < .01$.

5.3.4 Experiment 5. After reverse-coding appropriate items, an index of reactance was created by summing responses (see Dowd et al., 1991), with higher scores indicating higher reactance ($\alpha = .70$). Participants exhibited similar reactant scores ($M = 65.18$ (out of 112 in total), $SD = 6.43$, range = 52 to 83) to the normative data provided by Dowd et al. (1991; $M = 68.86$ (out of 112 in total), $SD = 6.59$, range = 46 to 83). A two-way ANOVA with goal source (internal vs. external) and goal–behaviour status (goal–behaviour discrepancy vs. goal–behaviour match) as between-participants factors for reactance was conducted. A significant main effect of goal–behaviour status indicated that reactance was higher in the

goal–behaviour discrepancy condition (M = 66.64, SE = 0.88) than the goal–behaviour match condition (M = 63.66, SE = 0.89), $F(1, 98) = 5.64$, $p = .019$, $\eta^2_p = .05$. No additional main effects or interactions were significant, all $p > .39$.

Appendix G: Systematic Analysis of Stereotype Accessibility Across Experiments 2-5

I analysed equivalent conditions across Experiments 2–5, grouping the large goal–behaviour discrepancy manipulation in Experiments 2 and 5 together, and the small goal–behaviour discrepancy manipulations in Experiments 3–5 together. The data were submitted to a 2 (Goal Source: internal vs. external) × 2 (Goal–Behaviour Status: goal–behaviour discrepancy vs. goal–behaviour match) × 2 (Word Type: Black-stereotypic vs. stereotype-neutral) × 2 (Word Valence: negative vs. positive) × 2 (Block: 1 vs. 2) mixed-model ANCOVA with goal source and goal–behaviour status as between-participants factors and Experiment number (i.e., Experiment 2 (78%), Experiment 3 (98%), Experiment 4 (100%), and Experiment 5 (78% and 100%) as a covariate to control for multiple experiments.

The analysis revealed four significant main effects. First, a significant main effect of block, $F(1, 352) = 10.44$, $p < .001$, $\eta^2_p = .03$, indicated that participants responded faster in Block 2 (M = 498 ms, SE = 3.19) than Block 1 (M = 526 ms, SE = 3.84). Second, a significant main effect of word type, $F(1, 352) = 13.29$, $p < .001$, $\eta^2_p = .04$, indicated that participants responded faster to Black-stereotypic words (M = 509 ms, SE = 3.36) than stereotype-neutral words (M = 515 ms, SE = 3.43). Third, a significant main effect of word valence, $F(1, 352) = 32.08$, $p < .001$, $\eta^2_p = .08$, indicated that participants responded faster to positive words (M = 498 ms, SE = 5.35) than negative words (M = 519 ms, SE = 6.49). Finally, a significant main effect of goal–behaviour status, $F(1, 352) = 7.62$, $p = .006$, $\eta^2_p = .02$, indicated that participants responded faster in the goal–behaviour discrepancy condition (M = 502 ms, SE = 4.99) than the goal–behaviour match condition (M = 522 ms, SE = 4.77).

The analysis also revealed two significant interactions. First, a significant Word Type × Word Valence interaction, $F(1, 352) = 7.01$, $p = .008$, $\eta^2_p = .02$. For positive words, participants responded faster to stereotype-neutral (M = 498 ms, SE = 3.19) than Black-

stereotypic words (M = 504 ms, SE = 3.53), $t(356) = 2.92$, $p = .004$, $d = 0.08$. For negative words, participants responded faster to Black-stereotypic words (M = 516 ms, SE = 3.56) than stereotype-neutral words (M = 532, SE = 4.00), $t(356) = 7.10$, $p < .001$, $d = 0.23$.

Second, and more importantly, a significant Goal Source × Word Type interaction, $F(1, 352) = 4.55$, $p = .034$, $\eta^2_p = .01$. Interaction means are presented in Figure G1. Participants primed with an internal goal responded equally fast to Black-stereotypic and stereotype-neutral words, $t(179) = 1.12$, $p = .26$, $d = 0.03$. In contrast, participants primed with an external goal responded faster to Black-stereotypic words than stereotype-neutral words, $t(176) = 3.86$, $p < .001$, $d = 0.12$. No additional theoretically relevant main effects or interactions were significant[43].

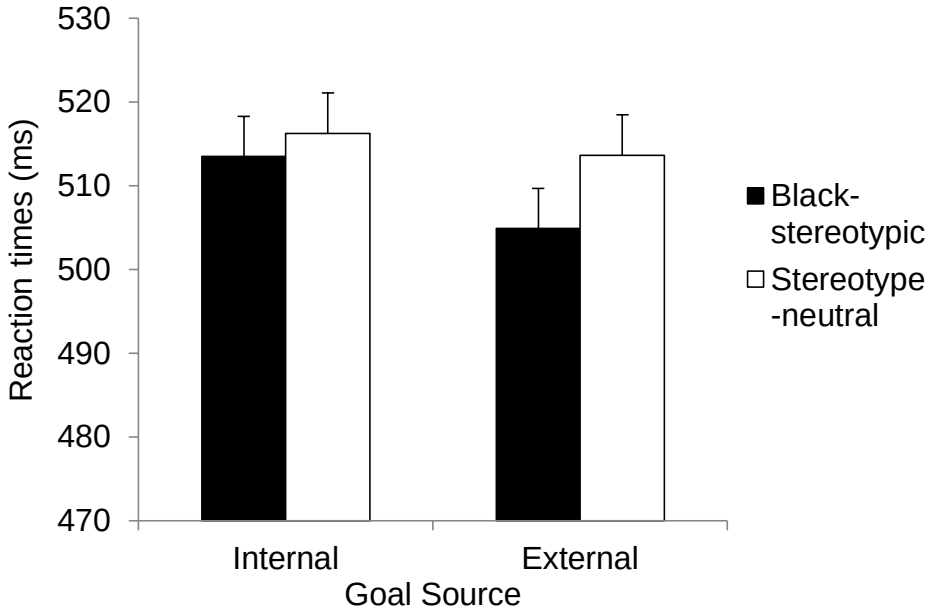

Figure G1. Mean reaction times (ms) as a function of goal source and word type, Experiments 2–5. Note. Error bars represent standard error.

[43] There was also a significant Word Type × Experiment interaction, $F(1, 352) = 6.07$, $p = .014$, $\eta^2_p = .02$. As this interaction is not pertinent to the current investigation, further analysis was not conducted. No additional main effects or interactions were significant, $p > .075$.