EFFECT OF GLUCOSE ON THE SUPPRESSION AND POST-SUPPRESSION
REBOUND OF STEREOTYPES

A thesis

submitted in fulfilment

of the requirement for the Degree

of

Master of Arts in Psychology

at the

University of Canterbury

by

Ryan John Paul Wilkinson

_____

University of Canterbury

July 2011

Table of Contents

List of Tables and Figures

Acknowledgements

Firstly, I would like to acknowledge and thank my supervisor, Professor Lucy Johnston, for all of her contributions to this thesis. Her guidance, extensive knowledge, and extreme patience have been invaluable and are appreciated. Also, thanks to my co-supervisor Professor Garth Fletcher for his advice and suggestions regarding regression analysis.

I am also thankful to my friends and family, particularly my mother and brother, for their continued support and encouragement over the past few years. Special mention must be made to my father, Ivan Wilkinson, who passed away while this research project was being undertaken. Special thanks also to Phillipa Wilding for all her input, support, and patience, especially during my thesis write-up.

Finally, big thanks to the participants of this study, as well as my fellow PG students and the staff at the Department of Psychology.

Abstract

The suppression of unwanted thoughts is an effortful process. An ironic effect of this process is that the unwanted thoughts can become hyper-accessibility after a period of their suppression, known as "post-suppression rebound". In the present study the impact of providing energy (through a glucose drink) on post-suppression rebound was investigated. One hundred and twenty participants participated in the main study, and another 30 participants served as a baseline group. Half of the participants in the main study were given a drink containing glucose and the other half was given a placebo drink containing an artificial sweetener. All participants wrote a passage about a "day in the life" of a gay male, with half the participants directed to avoid using stereotypes. A subsequent lexical decision task measured activation of stereotypes. Finally, a measure of prejudice was given to account for individual differences. Neither the direction to avoid using stereotypes nor the glucose resulted in lower stereotypicality of the "day in life" passages. Furthermore, response times during the lexical decision task did not differ between any of the main conditions or the baseline condition. However, the combination of both glucose and directed suppression did result in more positive passages, suggesting that the combination assists in reducing negative stereotype usage. Results are discussed in terms of stereotype usage and suppression and prejudice level.

Section 1: Introduction

The main aim of this thesis was to replicate and extend a previous study (Galliot, Peruche, Plant, & Baumeister, 2009) that found that giving participants glucose (as opposed to an artificial sweetener) resulted in fewer stereotypes being employed when writing about a gay male. The current study expands on these findings by exploring the effects of glucose on post-suppression rebound. This section will discuss the relevant theories, models, and previous research relating to the current study, prior to discussing the current study. . This section starts by describing what stereotypes are and why improving our understanding of when and why they are used can be of potential benefit to reducing the use of stereotypes and improving social interactions. Next, the section will examine what factors and cognitive mechanics are involved in the act of thought suppression, and why the suppression of thoughts may result in a greater activation of those thoughts later on. The section then focuses on previous research that has examined the post-suppression rebound effect with relation to the use and activation of stereotypes. Afterwards, the role that glucose (as a limited cognitive resource) has on tasks involving self-regulation (such as thought suppression) is explored. Lastly, the specific aims and hypotheses relating to the current study are detailed.

1.1 Stereotypes

A stereotype is a belief, or set of beliefs, about members of a specific social group, or type, of people. These beliefs are simplified and standardized concepts about the group members based on prior assumptions (which often contain a kernel of truth). A stereotype can exist based on any common element of a group of people, including (but not limited to) race, sex, religion, sexual orientation, or even hair colour. While

stereotypes are not inherently negative (for example "Asians are good at maths" and "Kenyans are good runners"), the use, expression, or endorsement of stereotypes can have a variety of negative impacts on those being stereotyped and can also negatively influence social interactions between the parties using stereotypes and members of the stigmatized group.

The use of stereotypes, or even the perception of the usage of stereotypes, can have negative impacts for both the perceiver and the target in a wide range of settings. For example, in male dominated areas such as engineering, academic records of female applicants are rated less favourably than similar records of male applicants (Foschi, Lai, & Sigerson, 1994). In criminal cases, African American defendants with more stereotypic facial features are likely to receive harsher sentences and are more likely to receive the death penalty (Blair, Judd, & Chapleau, 2004; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006).

Concerns over being stereotyped can similarly have negative consequences. Members of a stereotyped group may inadvertently reinforce stereotypes, as concern over conforming to social stereotypes can lead to a decrease in performance which is consistent with that stereotype, known as "Stereotype Threat" (Steele & Aronson, 1995; Shih, Pittinsky, & Ambady, 1999; Martens, Johns, Greenberg & Schimel, 2006). People who complain about being discriminated against are regarded less favourably by third parties, irrespective of whether their claims are justified (Kaiser & Miller, 2001).

The negative effects are not just exclusive to those being stereotyped. The perception of stereotype use can have negative impacts on the overall interaction. In order to protect their self-esteem, people will use the possibility of prejudice to devalue negative (though accurate) feedback if they can justify it in terms of the other

person being prejudiced (Crocker, Voelkl, Testa, & Major, 1991).). Concern over appearing prejudiced can also led to negative consequences for those who are attempting to regulate their behaviour. During inter-racial interactions, black participants rated high-prejudice white participants more favourably and as more engaged in the conversation than low-prejudice participants. It could be that high-prejudice participants were especially attentive in order to appear less prejudiced whereas low-prejudice participants were more relaxed and did not feel obliged to be attentive in order to appear less prejudiced. However, an alternative explanation is the low-prejudice participants were more focused on self-regulating their behaviour to appear less prejudiced than in engaging the other participant (Shelton, Richeson, Salvatore, & Trawalter, 2005). The latter explanation is in line with Vorauer and Taupie (2004), who found that during inter-racial interactions low-prejudice participants were more concerned with how they would be evaluated by their partner, and higher evaluative concern was associated with fewer intimacy-building behaviours.

While stereotypes are linked to prejudice and discrimination, knowing the first does not necessitate the others. It may be that simple knowledge of stereotypes can trigger their activation when exposed to a member of a stigmatized group; however personal beliefs, self-control, and cognitive effort may be able to inhibit the expression of these stereotypes (Devine, 1989; Lepore & Brown, 1997). An alternative theory suggests that mere exposure to a member of a stigmatized group and awareness of stereotypes may not result in their activation if the perceiver does not endorse those stereotypes (Lepore & Brown, 1997).

Despite the negative impact that stereotype use can have on the target, using stereotypes can be advantageous to the perceiver. By assuming traits and

characteristics of an individual based on their group membership, stereotypes can serve as energy-saving cognitive devices, simplifying social perceptions. The reduction in cognitive resources afforded by stereotyping was demonstrated by three experiments conducted by Macrae, Milne, and Bodenhausen (1994), in which participants had to concurrently complete a primary impression formation task and a secondary cognitive task. The results of these experiments showed that when participants were encouraged to use stereotypes they recalled more traits (both stereotypic and neutral) about the subject, and performed better on secondary tasks. It was suggested that being able to cognitively group a number of traits under one label frees up cognitive resources which can be used to attend to the neutral traits and the secondary task.

The relationship between cognitive load and stereotype use has also been demonstrated in other studies. When participants are required to make complex judgments on the guilt and aggression of a person, they tend rely on racial stereotypes and recall information consistent with those stereotypes. However, when those judgments are simplified, racial stereotypes are no longer a factor (Bodenhausen & Lichtenstein, 1987). When cognitive resources are taxed, self-regulation is poor, or motivation is low, the use of stereotypes allows for the preservation of cognitive resources. As such, the use of stereotypes is more likely in cases where there is a lack of ability or when there is no motivation, either internal (such as personal beliefs) or external (such as social pressures), to inhibit stereotypic thoughts (Macrae et al., 1994; Gailliot et al., 2007). Furthermore, using stereotypes can be beneficial to the perceiver, but by trying to avoid using stereotypes the perceiver may intentionally become more inclined to use them. This ironic effect of increasing post-suppression stereotype usage is due to the cognitive mechanisms involved in the suppression of

stereotypes. To explain this increased inclination to use stereotypes after suppressing them it is necessary to explain how thought suppression is proposed to work.

1.2 Thought Suppression

As stated above, trying to inhibit unwanted thoughts, such as stereotypes, may not be all that beneficial to the reduction of stereotype usage in the long-term. The act of thought suppression is not easily achieved. In a classic study investigating thought suppression, while completing two counter-balanced tasks participants were asked during one task to think about white bears (expression), and in other, to avoid thinking about white bears (suppression). Although suppression did reduce thoughts of white bears compared with expression, it did not eliminate them. Furthermore, during the expression task, participants indicated that thoughts of white bears were significantly higher when they had previously been instructed to suppress thoughts of white bears compared with those who had not suppressed previously (Wegner, Schneider, Carter, & White, 1987). Similar results have been found in replication studies (for example, Kelly & Kahn, 1994) as well as suppression of other stimuli, such as emotional or traumatic memories (Dalgleish, Yiend, Schweizer, & Dunn, 2009; Wenzlaff, Wegner, & Rober, 1988). These findings suggest that complete suppression is difficult to achieve and that the act of suppression can elicit greater subsequent expression of those target thoughts. The latter is referred to as "post-suppression rebound". By simply trying not to think about stereotypes a person may find that they cannot exclude all stereotypic thoughts; therefore once that effort is relaxed or can no longer be maintained, stereotype thoughts become more prevalent.

According to the model of thought suppression proposed by Wegner (Wegner & Erber, 1992; Wegner, 1994) the act of thought suppression requires two cognitive

processes of mental control: a controlled operating process and an automatic monitoring process, which work in a cyclical manner. The model is based on the premise that the more active a thought is, the more likely it is to influence subsequent thoughts and actions. In addition, thoughts require a certain threshold level of activation before they reach the person's awareness; however thoughts below that threshold level of activation can still influence behaviour (Devine, 1989; Macrae et al., 1994; Wegner & Erber, 1992; Wegner, 1994).

In Wegner's model, the function of the controlled operating process during thought suppression is to search for distracters, which serve to occupy the mind in place of the unwanted thought. The automatic monitoring process is proposed to search for suppression failures, which are occurrences where the unwanted thought is encroaching on awareness. However, the process of searching for a thought increases that thought's activation level. In order for the monitoring process to work it requires the unwanted thoughts to be cognitively accessible, in spite of the intended suppression, so as to seek out these encroaching thoughts. If unwanted thoughts are detected, they are replaced by distracters through the implementation of the operating process. Upon commencement of thought suppression both processes are activated. The monitoring process immediately identifies that the unwanted thoughts are in consciousness and activates the operating process. The operating process selects a distracter to absorb attention and then effectively turns off until the monitoring process recognises an association with current thoughts and the unwanted thought, and reactivates the operating process (Wegner & Erber, 1992; Wegner 1994).

According to Wegner, while the automatic process is relatively effortless, the controlled operating process requires some cognitive effort. As such, it is subject to the demands of finite attention resources. An implication of this model is that the

automatic process is constantly sensitising, or priming, the perceiver to the unwanted thoughts in order to monitor for their occurrence in awareness. When the suppression process is relaxed or the operating component fails, the priming of the unwanted thoughts (due to the monitoring process) leaves these thoughts hyper-accessible to the perceiver, resulting in post-suppression rebound. Motivation is also a key factor here. As the suppression process is cognitively effortful, even with adequate cognitive resources available, perceivers may not suppress their thoughts if not motivated to do so. Conversely, perceivers who are highly motivated to suppress certain thoughts may try not to relax the suppression process, and they are likely to become more resource efficient at suppressing thoughts. As a result of the constant suppression and reduced activation, highly motivated suppressors are less inclined to show a rebound effect (Kelly & Kahn, 1994; Wegner & Erber, 1992; Wegner 1994).

1.3 Post-Suppression Rebound of Stereotypes

As with other forms of thought suppression, consistent with Wegner's model, instructing participants to avoid using stereotypes during an initial task has been shown to increase activation of stereotypic thoughts. This increase in activation can be manifested in various ways, such as: an increase in usage of stereotypes following another writing task, altering behaviour to distance themselves from the target (for example, a skinhead) due to perceived negative stereotypes (hostility), and a hyper-accessibility of stereotype related words (Macrae, Bodenhausen, Milne, & Jetten ,1994).

The third experiment by Macrae et al. (1994) provided a model for the current study. In Macrae et al.'s experiment, participants spent five minutes writing about the "day in the life" of a skinhead, after which they completed a computer-based Lexical

Decision Task (LDT) that involved categorising letter strings as either words or non-words. Half of the word letter strings were considered stereotypic of skinheads and half were distracter words unrelated to skinheads. Results compared the response times of the stereotypic words to distracter words and showed that suppressers were faster to categorise stereotypic words compared with non-suppressers. These results are consistent with Wegner's model of thought suppression, and suggest that suppression of stereotypes may be counterproductive in the long term, eliciting higher stereotype use once active suppression ceases.

Studies have ruled out the rebound effect as being merely an artefact of directed suppression or from participants inferring that stereotypes are expected in subsequent takes by making use of "spontaneous suppression". Spontaneous suppression refers to when participants reduce their use of stereotypes without being explicitly directed to do so, and is largely dependent on the powerful influence of social norms. People can be highly motivated to conform to social norms in order to obtain a more accurate perception of reality, to form and maintain positive social interactions, and to maintain a positive self-image (Cialdini & Goldstein, 2004). Those who comply with social norms tend to be perceived more favourably by others (Shapiro & Neuberg, 2008). Whether it is due to the explicit expression of stereotypes generally becoming increasingly less common, the desire to maintain an egalitarian self-image, or the risk of social disapproval and negative sanctions on those who engage in prejudicial behaviour, social norms regarding stereotype use have been shown to provide strong motivation to suppress the expression of stereotypes (Castelli, Vanzetto, Sherman, & Arcuri, 2001; Crandall, Eshleman, & O'Brien, 2002; Shapiro & Neuberg, 2008).

As with instructed suppression, spontaneous suppression of stereotypes can also result in post-suppression rebound once the suppression is relaxed. Macrae, Bodenhausen, and Milne (1998) were able to get participants to spontaneously suppress stereotypes by manipulating self-focus which was accomplished via displaying either the participant's own image or that of a stranger on a television monitor. The belief was that high self-focus increases the salience of normative beliefs (such as avoiding the use of stereotypes) and promotes self-regulation. Adding support to the belief, participants used fewer stereotypes in high self-focus condition (own image displayed) than the low self-focus conditions (stranger's image). However, once the self-focus was reduced participants demonstrated a post-suppression rebound, increasing stereotype use compared with participants who always had a stranger's image displayed, or those whose own image was continually displayed.

The rebound effect following spontaneous suppression has been shown to be similar to the rebound following directed suppression. Wyer, Sherman, and Stroessner (1998) compared the rebound effect of directed suppression and spontaneous suppression conditions. Participants completed a survey on how much they endorsed various African American stereotypes before completing an impression formation task regarding a story about a person named "Donald" whose race is unspecified and who engages in ambiguously hostile activities. The premise was that the more a person is primed with thoughts about African American stereotypes the more hostile they would rate Donald to be. The experimenters achieved spontaneous suppression by telling a third of the participants that the survey was being conducted by an African American political group. A third of the participants were asked to avoid thinking

about stereotypes (directed suppression), and the remaining control participants were not given any additional information.

The results from Wyer et al. (1998) found that the participants in both the spontaneous suppression and directed suppression conditions rated Donald as significantly more hostile than those in the control condition, and that the spontaneous and directed suppression conditions did not differ significantly in the ratings of hostility. This suggests that both directed suppression and spontaneous suppression primed the participants, leading to a similar post-suppression rebound effect. Furthermore, the findings of Macrae et al. (1998) and Wyer et al. (1998) are consistent with Wegner's model of thought suppression. The fact that participants were not directed to avoid using stereotypes, suggests that it is the act of suppression itself that results in the rebound effect, rather than participants perceiving that stereotypes are expected in subsequent tasks or an artefact of directed suppression.

However, the results from studies on post-suppression rebound have not shown a rebound in every case. Monteith, Spicer, and Tooman (1998) reported that only high-prejudice participants showed post-suppression rebound of stereotype accessibility (reflected as a higher recall of stereotypic words than non-stereotypic words) while low-prejudice participants showed no rebound effect. This could suggest that stereotypic thoughts simply do not enter the minds of people who do not endorse the use of stereotypes (low-prejudice), consistent with findings such as those from Lepore and Brown (1997) who showed that while low-prejudice individuals were as aware as high-prejudice individuals of racial stereotypes they were less likely to spontaneous activate these stereotypes in response to the racial category.

An alternative suggestion is that low-prejudice participants may be more typically inclined to suppress stereotypic thoughts and as a consequence of repeated

practice at such, are more cognitively efficient at doing so; thus avoiding the increase in stereotype accessibility (Devine, 1989). While the present study did not directly compare these alternate points of view, the impact of prejudice level and availability of cognitive resources were considered and hence the findings might speak to the ability of each model to account for suppression effects. If low-prejudice participants have consistently low levels of stereotype use and an absent rebound effect, irrespective of cognitive resources and task instructions, then such results would tend to support the theory that stereotypes simply do not enter the mind of low-prejudice participants and are therefore less activated. Alternatively, if manipulations of cognitive resources results in variation of stereotype use and rebound among low-prejudice participants then this would be more in line with the idea that low-prejudice participants are more efficient at suppressing stereotypes, and increasing cognitive resources increases their efficiency.

As mentioned previously, thought suppression is an effortful act of mental self-regulation and thus requires self-control. Muraven, Tice, and Baumeister (1998) reported that participants who had previously engaged in a thought-suppression exercise were quicker to give up their attempt to solve an unsolvable problem, and were also impaired in their ability to control emotional expressions. It was concluded that the effortful nature of thought-suppression drains limited cognitive resources. The more time someone engages in self-regulatory acts, and the more effortful these acts are, the more these resources are depleted. This reduction in resources means that fewer resources are available for subsequent cognitive tasks, and as such performance is diminished. The reduction of cognitive resources through self-regulation is referred to as "regulatory depletion".

If the act of thought suppression both requires and depletes limited cognitive resources, increasing those resources should improve a person's ability to suppress thoughts. One such resource that has been investigated is glucose.

1.4 Glucose and Self-Control.

Glucose is the principle source of energy for the brain, and as the brain cannot store glucose it requires a constant supply from the bloodstream. Hypoglycaemia (low blood glucose levels) is associated with impaired cognitive functioning as well as other psychological and physiological problems. Recent studies have investigated the role that glucose, as a limited resource, has on mental effort, such as self-control and thought suppression. Blood glucose levels have been shown to deplete faster during tasks requiring more mental effort (e.g. incongruent Stroop task) than less demanding tasks (e.g. congruent Stroop task; Fairclough & Houston, 2004).

Similarly, tasks requiring self-control resulted in blood glucose depletion, which in turn predicted poorer performance on subsequent self-control tasks. In a task including an interracial interaction involving racially sensitive material, once initial measures of blood glucose were measured, participants (all white) engaged in a conversation about their opinions on affirmative action and criminal profiling with the experimenter (who was either white or black). Blood glucose measures were then taken a second time. The Internal Motivation to Respond Without Prejudice Scale (IMS; Plant and Devine 1998; as cited in Gailliot et al., 2007), was completed. This contains items such as "Because of my personal values, I believe that using stereotypes about Blacks is wrong". Participants who scored lower on the IMS (i.e., high-prejudice individuals) showed a greater depletion in blood glucose levels when conversing with a black experimenter than with a white experimenter. Participants

who scored higher on the IMS (low-prejudice) had the reverse effect, with blood glucose levels depleting during conversations with white experimenters, though to a lesser extent than the high-prejudice participants conversing with the black experimenter. The authors suggested that participants who had low IMS scores (high-prejudice) were not typically inclined to stifle prejudicial thoughts and accordingly it took more effort to do so in this task. Whether it is their usual tendency to not suppress prejudicial thoughts or simply their endorsement of stereotypes and prejudicial thoughts, the results suggested the low IMS participants (high-prejudice) required higher self-control to navigate through a racially sensitive topic with a member of a different race, under the constraints of social norms. This higher demand for self-control relates to a higher demand from the brain for glucose; thus glucose in the blood is diverted to the brain, depleting the levels of blood glucose (Gailliot et al., 2007).

By ingesting a glucose-rich drink between self-control tasks, blood glucose levels can be restored and the subsequent decrease in performance can be avoided. In Galliot et al.'s (2007) seventh study, after an initial Stroop task (to establish a baseline), participants engaged in an attention control task. Half the participants simply watched a video and the other half were instructed to focus on the woman in the video and ignore the common single-syllable words also displayed (a task that had previously been shown to deplete blood glucose levels). Participants were then given a lemon drink sweetened with either sugar (sucrose; a combination of fructose and glucose) or an artificial sweetener (sucralose). In order to allow for any glucose to metabolise, twelve minutes were allowed to pass before the participants completed a second Stroop task. Results showed that participants who were given the glucose drink performed equally well on the second Stroop task regardless of which attention

control condition they were in. Participants who were given the artificial sweetener showed a significantly higher number of errors during the second Stroop task if they were in the controlled attention condition than if they just watched the video normally. These results support the idea that restoring blood glucose levels after a self-control task restores cognitive performance. Similar findings were obtained when participants were given the lemon drink, and then instructed to write about the emotions invoked with regards to either the thought of their own death (mortality salience condition) or dental pain, before completing a crossword puzzle. The mortality salience with placebo drink condition differed from the other three, with participants leaving more words unsolved on the crossword puzzle, which was the dependent measure of persistence and self-control (Gailliot et al., 2007). This finding suggested that the extra glucose aids to protect against the glucose-draining effects such as those associated with the thoughts of one's own death.

In a later study by Galliot et al. (2009), participants were again given a drink containing either sugar or an artificial sweetener. Following twelve minutes of completing filler questionnaires, participants were shown a picture of a young man called "Sammy" who was said to be gay. Sammy served as the stimulus for a "day in the life" task. Participants were not instructed to suppress or avoid using stereotypes, yet those participants who had received the extra glucose by consuming the sugar drink used fewer stereotypes than those in the control condition (artificial sweetener). As expected, participants who scored higher on explicit prejudice used more stereotypes when writing about Sammy. The reduction in stereotype usage in the glucose condition did not differ significantly between low and high prejudice participants; however high-prejudice scoring participants used fewer prejudicial statements in the glucose condition than in the control condition. Drink condition had

no effect on low-prejudice participants. This suggests the extra energy provided by the glucose better enabled participants to spontaneously suppress their stereotypes (in accordance with social norms) due to the increase in resources required for self-control and, related to the experiment on interracial interactions (Gailliot et al., 2007), allowed participants high in prejudice to suppress prejudicial thoughts.

The above study by Gailliot et al. (2009) demonstrated that by ingesting glucose people are more likely to spontaneously suppress their use of stereotypes. An implication of this study could be that maintaining a sufficiently high blood sugar level could help alleviate the use of stereotypes when interacting with a stereotyped group. However, the study did not address the effect that sucrose may have on any post-suppression rebound. The increase in blood glucose may assist participants in spontaneously reducing their explicit use of stereotypes, however if the glucose is helping to fuel the engine driving thought suppression then, according to Wegner's model, the automatic process should still be priming the participant with stereotypic thoughts. When the social requirement for suppression is relaxed, an increase in stereotype activation may occur.

The aim of the current study was to measure stereotype activation following glucose-assisted spontaneous suppression and investigate whether there were differences between a glucose-assisted spontaneous suppression with that of a directed suppression (with both a glucose condition and a placebo condition).

 1.5 The Current Study

The current study aimed to replicate the results of Gailliot et al. (2009) and to investigate the possible occurrence of post-suppression rebound after suppression of stereotypes when glucose has or has not been consumed. As such, this study was

based largely on the experiment conducted by Gailliot et al. (2009), and the third experiment conducted by Macrae et al. (1994). As in the study by Gailliot et al. (2009) participants in the main study consumed a drink containing either sugar (glucose condition) or an artificial sweetener (placebo condition) prior to completing a "day in the life" passage about a gay male. Extending on the previous study, and following the study by Macrae et al. (1994), half the participants within each drink condition were instructed to avoid using stereotypes (directed-suppression condition) during the passage while the other half were given no such instruction (control condition). Participants subsequently completed a Lexical Decision Task (LDT) to measure stereotype activation. In this study, half of the words were stereotypic of gay males, and the other half were distracter words. Participants lastly completed a measure of explicit attitudes towards homosexuals (Heterosexuals Attitudes Towards Homosexuals scale). Another set of participants completed only the LDT to provide baseline stereotype activation data

As all participants in the main study would be primed with thoughts of a gay male during the "day in the life" passage, it was expected that all of these participants would show a faster response time to words associated with gay males than to the distracter words. Since participants in the baseline condition would not have been primed with the information of a gay male (Sammy), they should show no difference between gay stereotype words and distracter words.

The main hypothesis of this study is that the glucose will assist in suppression of stereotypes in the "day in the life" passage, and as such will result in similar, high levels of post-suppression rebound in the LDT (demonstrated by faster response times to stereotype words compared with non-stereotype words), irrespective of the whether participants were directed to suppress stereotypes or not. The reason for the

hypothesised similarity in rebound effect between passage instructions is that participants who ingest sucrose in the no instruction condition are expected to spontaneously suppress stereotypes, as was reported in Gailliot et al. (2009).

Participants in the directed suppression condition who consumed the artificial sweetener were also expected to suppress stereotypes during the writing task and thus show a post-suppression rebound, although response times to both sets of words were expected to be slower than the other suppression conditions due to glucose assisting the overall performance in the other conditions. Participants in the control condition who consumed the artificial sweetener and who were not instructed to suppress stereotypes were expected to use the most stereotypes in the "day in the life" passage, and to show the least difference in response time between the stereotype and distracter words (i.e. less suppression therefore less rebound). However, some suppression may occur due to prevalent social norms against stereotyping, therefore participants in the control condition were expected to show a greater difference in response times to word type than in the baseline condition. The rationale for the higher rebound effect in the other conditions compared with the placebo-control condition was that glucose should be beneficial during self-control tasks ("day in the life"), but the increased ability to suppress stereotypes will subsequently result in a higher accessibility of those stereotypes during non-self-control based tasks such as the LDT (in accordance with Wegner's model of thought suppression).

Prejudice is another factor likely to influence results, with a negative correlation between Heterosexual Attitudes Towards Homosexuals (HATH; Larsen, Reed, & Hoffman, 1980) scores and stereotype use in the "day in the life" passage task. That is, those with a higher HATH score (low-prejudice) would show greater stereotype use and a reduced rebound effect consistent with Monteith et al. (1998). As

was found in Gailliot et al. (2009), there was not expected to be a significant

interaction between HATH and sucrose on stereotype use or accessibility. While no

sex differences were reported in either Gailliot et al. (2009) or Macrae et al. (1994),

being male is a predictive factor of a low score on the HATH scale (Larsen et al.,

1980) and as such may be factor in stereotype use, though other studies have reported

no sex-difference in stereotype beliefs of homosexuals (LaMar & Kite, 1998).

In summary, it was hypothesised that participants in the placebo-control

condition (no suppression) would use more stereotypes than those in the other

(suppression) conditions during the "day in the life" task. Secondly, all of those who

previously suppressed their stereotypes in the "day in the life" task were predicted to

show a greater difference in response time between the stereotype-words and the

distracter-words (post-suppression rebound) during the LDT, compared with those in

the placebo-control condition. Those in the baseline condition were expected to show

the least difference between word-types.

Section 2: Method

<u>2.1 Participants</u>

One hundred and twenty participants (60 males, 60 females) volunteered to participate in the main study. The average age of the participants was 21.6 years (SD = 6.4). Of the 120 participants, 112 classified their sexual preference as heterosexual, two participants identified themselves as bisexual, one as homosexual, one participant categorised their sexual preference as "other", and four preferred not disclose their sexual preference. One hundred of the participants were recruited from the University of Canterbury 100-level psychology participant pool in which students participated in return for course credit. The remaining 20 participants were recruited through email and notice board advertisements, and were given a $10 Motor Trade Association (MTA) voucher for their time. Participants were randomly assigned to one of the four conditions, with 30 participants allocated to each condition. To ensure equal numbers across conditions the random assignment based on order of the participants was generated prior to running the study. However, the predetermined assignment resulted in an unequal number of males and females in some conditions.

An additional 30 participants (23 females, 7 males) were recruited as a baseline comparison group for one of the experimental tasks (the LDT – see below). These participants were recruited separately via emails and notice board advertisements, and were given a $5 cafe voucher for their time. The average age for the baseline participants was 23.8 years (SD = 7.1).

The study used a 2 (Beverage Sweetener: Glucose or Artificial Sweetener) x 2 (Essay instruction: Suppression or Control) x 3 (LDT Word Type: Stereotype, Distracter, or Non-Word) mixed design with repeated measures on the third factor. A baseline condition (LDT word type only) was included for comparison.

<u>2.2 Materials</u>

Day in the Life

A black and white picture of a young man called Sammy was used as the stimulus in the "day in the life" passage. A brief description was shown below the picture, which was claimed to be written by Sammy. In this description Sammy described himself as 23 years old, living and working in Wellington, gay, and enjoying hanging out with his mates and watching movies to relax. (See Appendix A)

Heterosexual Attitudes Towards Homosexuals (HATH) Scale

 The 20-item Heterosexual Attitudes Towards Homosexuals (HATH) scale (Larsen et al., 1980) was used to measure explicit attitudes to homosexuals. The scale uses a seven-point Likert scale (1= strongly disagree, 7 = strongly agree) with statements such as "I enjoy the company of homosexuals" and "It would be beneficial to society to recognize homosexuality as normal". Half of the statements are phrased in a negative manner (e.g. "Homosexuality is immoral") and require reverse scoring. Once the 10 items have been reverse-scored, the average across all 20 items is calculated for each participant. A higher score reflects a more positive explicit attitude towards homosexuals.

Beverages

The two drinks used were 400 millilitres of Baker Halls® Lemon and Barley fruit juice syrup (original and low-calorie versions) mixed in a ratio of appropriately 1:6 with water. The original version (glucose drink) contains sucrose (sugar) and has approximately 144 calories when mixed, whereas the low-calorie version (placebo drink) contains the artificial sweetener Sucralose and has approximately 18 calories

when mixed. This is a similar volume and calorie content of the drinks given in Gailliot et al. (2009). The drinks were mixed by an independent person and placed into two distinct containers. The researcher was unaware which container corresponded to which beverage condition, making both the researcher and participants blind to this condition.

Beverage and Demographics Questionnaire

A beverage questionnaire measured how much participants liked the beverage they were given ("How pleasant was it for you while drinking the beverage?", "Did you enjoy the taste of the beverage?", "Did you find the beverage refreshing?") using a seven-point Likert scale. Participants were asked to estimate the number of calories the beverage had (given that there are approximately 140 calories in an average soft drink). The purpose of these questions was to assess whether participants could tell which drink had low-calories and whether there was a preference of one of the drinks over the other. This questionnaire also asked for demographic information (sex and age), as well as asking if the participants had consumed anything (except plain water) in the past three hours, and if so, to detail what they had consumed. Three participants reported that they had consumed something in the past three hours (two had coffee, one had chewing gum). Exclusion of these participants' data did significantly alter the results and was therefore included in the analyses reported below.

Filler Questionnaire

A questionnaire on workplace scenarios was used to fill in time to allow any glucose to be absorbed into the blood stream before participating in the experimental tasks. This questionnaire was chosen because it would occupy the appropriate amount of

time, it was not too cognitively demanding, and it was not considered likely to have

any relevant priming effects.

Lexical Decision Task (LDT)

 A custom-written computer-based Lexical Decision Task (LDT; Walton, 2003) was

used to measure stereotype accessibility. For each trial a string of letters was

presented in the centre of the computer screen. Participants were required to

categorise each letter string as either a "word" or "non-word" by means of a key press

(L-key for "Word", A-key for "Non-word"). The list of letter strings were displayed

in a unique random order for each participant and consisted of 14 words stereotypic of

homosexual males, 14 matched distracter words, and 28 non-words (see Appendix B).

The 14 stereotypic words were determined via a pilot study in which 12 participants

(seven females, five males) from the same general population as the participants in the

main study rated each of a list of 57 traits and characteristics frequently associated

with stereotypes of gay males (Robinson, Montiel, Jakubowski, & Madon, 1996; as

well as experimenter selected traits), according to how much they associated those

traits with homosexual males and how much they associated those traits with

heterosexual males using a seven-point Likert scale . The 14 stereotype words chosen

for use in the main study were those traits that showed the highest mean ratings to

homosexual men (compared with heterosexual men; see Appendix C). Participants in

the pilot study also rated how positive or negative each trait was. The 14 matched

distracter words were selected from common words with similar structure and

identical letter length to the stereotypic words. The 28 non-words were matched to

the stereotypic words for word length, and were derived by changing a letter or two of

common words so that they remained pronounceable and followed the phonetic rules of the English language.

2.3 Procedure

Participants were tested individually. Prior to arriving, participants in the experimental conditions were randomly assigned to one of the four conditions (glucose with directed suppression, glucose without directed suppression, artificial sweetener with directed suppression, or artificial sweetener without directed suppression) and were requested to fast for three hours prior to the study. Upon arrival participants were briefed on the cover story, that the experiment was investigating the effect various food and beverages have on impression formation and cognition. Slight deception was required here as if participants were informed that experiment was investigating stereotype usage this may have affected behaviour due to social desirability or expectation bias. The lemon and barley drink containing either sucrose or the artificial sweetener was given to all of the participants. Participants then spent approximately 12 minutes completing a three-item questionnaire on the drink, a demographics questionnaire, and the filler questionnaire. This was to allow any sucrose to be broken down into glucose and absorbed into the blood stream (Gailliot et al., 2007; Gailliot et al., 2009).

After completing the questionnaires, participants were told that the next part of the study was measuring impression formation. They were told that they had been randomly assigned to consider one of 12 people who had provided a picture as well as a brief description of themselves. Actually, all participants were assigned the same person, Sammy, who described himself, amongst other things, as gay. Participants were then instructed to write an essay about a typical day of the man for five minutes.

Half the participants in each experimental condition were given additional instructions requesting that they try to avoid using stereotypes of any groups that the individual had identified him or herself as being a member of (directed suppression). The other half will be given no such instruction. To preserve their anonymity, participants were instructed to place their passages in a drop-box marked "day in the life" after five minutes. After the "day in the life" essay, participants then completed the LDT to measure stereotype activation. Participants were informed that the LDT was measuring cognitive aspects of performance. Participants were showed a series of letter strings one at a time on the computer screen. The participants' task was to indicate (by button press), as quickly and as accurately as possible, whether the string of letters was a word or non-word. Participants in the baseline condition completed only the LDT task.

After the LDT task participants in the experimental conditions were asked to complete the HATH scale. As with the "day in the Life" passage, a drop-box marked "HATH" was also provided. Finally, participants were debriefed regarding the actual study, asked to re-consent to their results being used, given the incentive (if not recruited from the participant pool), and thanked. Participants were told that it would be useful to know their sexual preference and if they felt comfortable doing so, they were invited to write down their sexuality (homosexual, heterosexual, bisexual, or other) which they placed in an envelope before leaving.

This research project was reviewed and approved by the University of Canterbury Human Ethics Committee.

2.4 Data Collation and Preparation

The dependent variables were the ratio of stereotype words over total words used in "day in the life" essay, stereotypic and positivity ratings of the essay, score on

the HATH scale, response time to distracter and stereotypic words (in milliseconds) during LDT, and calorie estimation of drink. Demographic information (age and sex) were acquired for reporting purposes and analyses (with sex a possible factor in stereotype use and HATH).

Two independent raters were hired to examine the "day in the life" passages. The passages were rated on a nine-point Likert scale according to how stereotypic of homosexuals they thought the content of each of the passages was (1= not all stereotypic, 9 = very stereotypic), and how positive they thought each of the passages was (1= very negative, 9 = very positive). Raters also coded each passage for any use of the 57 homosexual stereotype traits and characteristics (or synonyms) examined in the pilot study (based on the study by Robinson et al., 1996). Raters also counted the total number of words used in each passage and then divided the number of stereotypic traits by the total number of words to give a stereotype ratio. This gave two measures of the passage stereotypicality (stereotype rating and stereotype ratio), and a measure of how positive or negative the passages were (positivity rating). The inter-rater reliability of the each measures was below the .7 ideal with, r=.61 (stereotype rating), r =.62 (stereotype ratio), and r =.50 (positivity rating). Despite the low inter-rater reliability, a mean rating across the two raters was used. The appropriate items on the HATH scale were reverse scored and the mean rating on the scale was used as a measure of explicit prejudice (α = .925). Data cleaning procedures were applied to the response times of the LDT. Incorrect responses in the LDT were excluded from analysis of response times (n=1450, 17.3% of total responses). To ensure an approximate normal distribution a log transformation was applied to the response times. Response times that were three standard deviations above or below each participant's mean response time were also excluded to remove outliers (n=120,

1.73% of total correct responses). For the three letter string groups (stereotype words,

distracter words, and non-words) the mean response time for each participant was

then calculated.

Section 3: Results

3.1 Descriptive Statistics

Means and standard deviations for the dependent measures are shown in Table 1 as a function of experimental condition, with the exception of the response time (LDT task) and stereotypic measures ("day in the life" task) that are discussed below.

Table 1: Means (and standard deviations (SD)) of dependent measures as a function of condition

Measure	Condition					Overall
	Glucose-Directed	Glucose-Control	Placebo-Directed	Placebo-Control	Baseline	
Age	22.9	21.6	21.7	20.4	23.8	**22.1**
	(8.8)	(5.5)	(6.9)	(3.2)	(7.1)	**(6.6)**
# Males	13	14	18	15	7	**67**
# Females	17	16	12	15	23	**83**
HATH Score	5.9	6.0	5.7	6.0	-	**5.9**
	(1.0)	(0.9)	(0.7)	(1.0)		**(0.9)**
Beverage Rating	5.6	5.7	5.2	5.4	-	**5.5**
	(1.1)	(1.0)	(0.7)	(0.9)		**(1.0)**
Calorie Estimate	102.9	102.4	98.6	99.8	-	**100.9**
	(36.8)	(36.1)	(39.0)	(28.7)		**(35.0)**

The HATH score was not normally distributed with skewness of -1.10 (SE = 0.22) and an excess kurtosis of 0.73 (SE = 0.44), indicating a ceiling effect.

A 2 (Beverage) x 2 (Suppression) x 2 (Sex) ANOVA on explicit measure of prejudice (HATH) revealed only a significant main effect of sex, $F(1,112) = 9.47$, $p <$

.01, η_p^2 =.078, with males scoring lower on the HATH (higher prejudice), (M = 5.61, SD = 0.94) than females (M = 6.13, SD = 0.81).

3.2 Manipulation Checks

Beverage:

The participants' ratings on the three-item drink questionnaire had a high internal consistency (α = .822) and were averaged to obtain an overall beverage rating. Overall, the glucose drink was rated more positively (M = 5.64, SD = 1.03) than the placebo drink (M = 5.32, SD = 0.85), with the difference approaching statistical significance, *t* (118) = 1.90, *p* = 0.06, η_p^2 = .03

There was no significant difference in the mean estimated calorie content of the two drinks. The comparison of drink ratings and estimated calorie content suggests that although the glucose drink was rated more positively than the placebo drink, this was not due to a perceived difference in calories.

3.3. Preliminary Analysis

Sex:

The effect that the sex of participants had on each of the dependent variables was analysed and is shown in Appendix D. Analyses showed that there were significant differences between males and females with regards to HATH scores, positivity ratings, and LDT response times to all three word types. Sex will be used as a predictor variable only in further analysis of the dependent variables where sex has shown to have a significant effect.

Stereotypicality and Positivity:

The distribution of the stereotype ratio during the "day in the life" passage was not normally distributed, with a skewness of 2.28 (SE = 0.22) and excess kurtosis of 5.68 (SE = 0.44). Similarly the distribution of the total number of stereotypes used was not normally distributed with a skewness of 2.47 (SE = 0.22) and excess kurtosis of 7.08 (SE = 0.44). The skewness and kurtosis of these distributions indicate that both are highly peaked with floor effects, which may negatively affect further analysis and interpretation due to the assumption of normal distribution of variables.

The distribution of the stereotype rating had a skewness of 0.93 (SE = 0.22) and excess kurtosis of -0.23 (SE = 0.44), indicating a relatively normal distribution with a slight floor effect. The positivity rating was also relatively normally distributed with skewness of -0.31 (SE = 0.22) and an excess kurtosis of 0.92 (SE = 0.44). The relatively normal distribution of the stereotype rating and positivity ratings satisfies the assumption of normally distribution of data required in some analyses.

3.4. "Day in the Life" Essay

As shown in Table 2, participants in the conditions with directed suppression used fewer stereotypes across all of the stereotype measures than did participants in the control conditions. Participants in the glucose-control condition also used fewer stereotypes than the placebo-control condition.
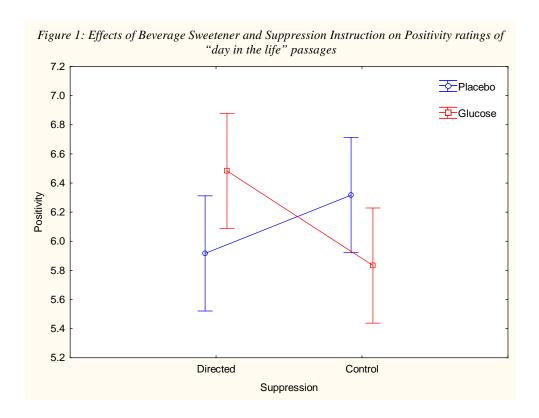
Table 2: Means (and standard deviations) for the measures of stereotypicality of the
"day in the life" passage as a function of condition

Measure			Condition		Overall
	Glucose-Directed	Glucose-Control	Placebo-Directed	Placebo-Control	
No. of Stereotypes	0.88 (1.44)	1.03 (1.66)	0.85 (1.25)	1.10 (1.28)	**0.97 (1.40)**
Stereotype Ratio	0.009 (0.013)	0.011 (0.017)	0.010 (0.015)	0.012 (0.015)	**0.011 (0.015)**
Stereotype Rating	3.17 (2.24)	3.30 (2.50)	2.65 (1.81)	3.85 (2.44)	**3.24 (2.28)**
Positive Rating	6.48 (0.90)	5.83 (1.25)	5.92 (1.13)	6.31 (1.06)	**6.14 (1.11)**

However, 2 (Suppression Instruction: Directed/Control) x 2 (Beverage:
Sucrose/Artificial sweetener) ANOVAs on both the stereotype ratio and the
stereotype rating measures revealed no significant effects. Further, regression analyses
with stereotype ratio and stereotype rating as dependent measures and beverage and
suppression conditions and HATH scores and sex as predictor variables revealed no
significant main effects or interactions using either measure of stereotype use. T-tests
comparing the placebo-control condition to the other three conditions using a
Bonferroni adjustment ($\alpha = .017$) on stereotype rating and stereotype ratio were not
significant.

For the positivity ratings, a 2 (Suppression Instruction) x 2 (Beverage) x 2
(Sex) ANOVA revealed a significant main effect of sex $F (1, 112) = 5.973$, $p < .05$,
$\eta_p^2 = .051$. Passages written by males (M = 5.88, SD = 1.02) were rated as being less

positive compared with the passages written by females (M = 6.39, SD = 1.15). The

two way interaction of Beverage x Suppression Instruction was also significant, F (1,

112) = 6.518, p < .05, η_p^2 = .048, as shown in Figure 1.

Figure 1: Effects of Beverage Sweetener and Suppression Instruction on Positivity ratings of "day in the life" passages

Post-hoc (Fisher's LSD, p <.05) tests compared effects of drink type

separately for each suppression condition and effect of suppression condition

separately for each type of drink. These tests revealed two significant differences:

between the drink types in the directed suppression conditions and between

suppression conditions for the glucose drink condition. When directed to suppress

stereotypes those in the glucose condition wrote more positive passages than those in

the placebo condition (Ms = 6.48 vs. 5.92). For the glucose drink, those who were

directed to suppress stereotypes wrote more positive passages than those in the control

condition (Ms= 6.48 vs. 5.83). The pair-wise comparison also showed that the placebo

control condition did not significantly differ from any of the other conditions.

3.5 Lexical Decision Task (LDT)

Mean response times for correct responses and error rates are shown in Table 3, as a function of condition and word type.

Table 3: Mean response times (excluding errors and outliers) in milliseconds and the mean number of errors made as a function of experimental condition. (Standard deviations shown in parentheses)

Measure	Condition					Overall
	Glucose-Directed	**Glucose-Control**	**Placebo-Directed**	**Placebo-Control**	**Baseline**	
Response Times (ms)						
Stereotype	798	808	795	746	752	**780**
	(282)	(249)	(185)	(126)	(172)	**(209)**
Distracter	842	803	847	807	786	**817**
	(305)	(222)	(223)	(194)	(146)	**(222)**
Non-Word	1008	1008	1072	966	971	**1005**
	(346)	(378)	(337)	(226)	(311)	**(322)**
Errors						
Stereotype	1.63	2.37	1.73	1.47	1.13	**1.67**
	(1.38)	(2.54)	(1.55)	(1.14)	(1.14)	**(1.66)**
Distracter	1.40	2.03	1.63	1.13	1.13	**1.47**
	(1.67)	(2.53)	(1.54)	(1.14)	(1.28)	**(1.71)**
Non-Word	7.57	6.30	6.06	7.13	5.47	**6.51**
	(4.26)	(3.61)	(3.61)	(4.38)	(3.41)	**(3.90)**

Error rates: A 2 (Beverage) x 2 (Suppression) x 2 (Sex) x 3 (Word Type) ANOVA, with repeated measures on the final factor, on the errors, showed only a main effect of word type, F (1.22, 136.36) = 193.025, p <.001, (using Greenhouse-Geisser correction due to violation of sphericity), η_p^2 = .633. Post-hoc (Fisher's LSD, p <.05) tests revealed the difference in errors due to word type was due to the

significantly higher errors in response to non-words (M= 6.77) compared with stereotype and distracter words (Ms= 1.8 and 1.55, respectively). The difference in errors between stereotypic words and distracter words was not significant.

Response times: As shown in Table 3, response times to non-words were slowest in all groups. In all but one of the conditions (glucose-control), response times to distracter words were next slowest, with response times to stereotype words being the fastest. In the glucose-control condition, distracter word response times were slightly faster than the stereotype response times.

Analysis of variance was carried out using the log of response times as the dependent variable, but means are reported in the text for ease of interpretation. A 2 (Beverage) x 2 (Suppression Instruction) x 2 (Sex) x 2 (Word Type[1]) ANOVA, with repeated measures on the final factor, revealed significant main effects of word type, $F (1, 112) = 12.86$, $p <.001$, $\eta_p^2 = .103$; and sex, $F (1, 112) = 6.22$, $p <.05$, $\eta_p^2 = .053$. Response times to stereotypic words (M = 786ms, SD = 217ms) were faster than to non-stereotypic (distracter) words (M = 825ms, SD = 237ms). Female participants responded faster (M = 761ms, SD = 210ms) than male participants (M = 851ms, SD = 237ms).

Separate single factor ANOVAs and t-tests were computed for the stereotypic and distracter words comparing the five LDT conditions (Glucose-Directed, Glucose-Control, Placebo-Directed, Placebo-Control, and Baseline). These showed that there was no significant difference between the five conditions for either the stereotype words or distracter words.

---

[1] Non-words were included in the LDT task only to prevent anticipatory "Yes" responses but are not considered in analysis as key comparison is between stereotypic and non-stereotypic words

3.6 Relationship Between Variables.

Pearson product-moment correlations using a Bonferroni adjustment ($\alpha = .0074$) were computed between the HATH score, response time, stereotypic measures, and beverage rating, and are shown in Table 4. The dependent variables in the "day in the life" passage (stereotype ratio, stereotypic rating, positivity rating) significantly correlated with each other. The response times during the LDT also correlated with each other, however there were no significant correlations between variable of different tasks or between the HATH score. The beverage rating correlated positively with the stereotypic rating indicating that the more participants preferred the drink the more stereotypic their "day in the life" passages were. The relationships between variables did not significantly differ between sexes.

Table 4: Pearson product-moment correlation matrix of dependent variables

		1	2	3	4	5	6
1	HATH Score						
2	Beverage Rating	-.05					
3	Stereotype Ratio	-.09	.17				
4	Stereotype Rating	-.13	.27	.75			
5	Positivity Rating	.11	-.01	.38	.32		
6	Stereotype Response Times	-.19	.03	.14	.10	-.23	
7	Distracter Response Times	-.21	.00	.14	.08	-.14	.87

Note: the correlations greater than or equal to $\pm.27$ are significant at $p < .007$

Section 4: Discussion

The aim of the current study was to replicate the results of Gailliot et al. (2009), that participants used fewer stereotypes when given a drink containing glucose, and to investigate the possible occurrence of post-suppression rebound after suppression of stereotypes when glucose had or had not been consumed. Firstly, it was hypothesised that participants who were not instructed to avoid using stereotypes would use fewer stereotypes when writing about the "day in the life" of a gay male when they had previously ingested a drink containing glucose compared with a placebo drink containing an artificial sweetener. It was also hypothesised that participants who were given drinks containing glucose and participants who were directed to avoid using stereotypes would show a greater activation of stereotypes, resulting in a faster response time to stereotype words during the subsequent lexical decision task.

The primary findings of the study indicated that there was no significant difference in the overall use of stereotypes between those who had ingested glucose and those who had not, nor was any difference or interaction found in the use of stereotypes between participants who had been directed to avoid using stereotypes and those who had not been given such a direction. However, an interaction effect did occur with regards to how positive the passages were, with participants writing more positive passages after having been given glucose and when directed to avoid using stereotypes compared with other conditions. This result suggests that the combination of glucose and directed suppression reduces the use of negative stereotypes. In addition, it was found that drink type and suppression instruction had no significant effect on the difference in response times between stereotypic and non-stereotypic words.

Each of these primary findings will be discussed in further detail in this section, including significant differences that were observed. The findings will also be compared with findings from previous research and current theoretical models. Methodological limitations and suggestions for future research will be discussed. In addition the possible implications and applications of this research will be explored. Finally, concluding comments will provide an overview of the results and findings detailed in this study

4.1 Suppression of Stereotypes

This study was largely based on that of Gailliot et al. (2009) which found that participants who were given a drink containing glucose used fewer stereotypes when writing about the "day in the life" of a gay male in comparison to participants who had been given a drink containing an artificial sweetener. The researchers proposed that their results illustrated the important role that self-regulatory resources, such as glucose, play in the suppression of stereotypes. The results of the current study do not directly support the findings of Gailliot et al (2009). The drink that the participants were given in this study did not have a significant effect on the overall use of stereotypes. These results are also not consistent with other findings such as those by Gailliot et al. (2007) and Fairclough and Houston (2004) that low blood glucose resulted in decreased performance during tasks that required mental effort and self regulation. The lack of effect in reducing the stereotypicality of the passages in the current study may be due to an overall low use of stereotypes, which is addressed in subsequent paragraphs.

Due to the influence of social norms, it was hypothesised that participants who were not directed to avoid using stereotypes would avoid using stereotypes anyway

resulting in a spontaneous suppression. Participants who were given glucose drinks were expected to be better equipped to regulate their behaviour and therefore show a greater degree of spontaneous suppression. This improved spontaneous suppression did not occur. However, it is worth noting that it was expected that participants who were directly asked to avoid using stereotypes during the "day in the life" essay would also use fewer stereotypes because they were directed to do so. While participants in the directed suppression conditions did use fewer stereotypes than those in the control suppression conditions this difference did not approach statistical significance. This result is not in accordance with other studies (Macrae et al.,1994; Monteith et al., 1998; Wyer, Sherman, & Stroessner, 2000; Zhang & Hunt, 2008), which found that similar directions to avoid using stereotypes during a "day in the life" passage resulted in fewer stereotypes being used. Overall, it was hypothesised that the placebo-control condition would use more stereotypes than the other three conditions; however these differences were not statistically significant.

The lack of statistical significance from the drink effect or suppression instruction with regards to the passage stereotypicality may in fact be due to the overall lack of stereotype usage across conditions. This is possibly due to the low prejudice of the sampled population or possibly due to strong social norms against stereotyping in all conditions. Of the passages written during this study, over a third made no reference to any of the stereotype traits and characteristics (or synonyms), irrespective of condition Furthermore, the total number of stereotypes used and the stereotype ratio distributions showed a floor effect with little variance. The mean number of stereotypes used in other studies (Gailliot et al., 2009; Monteith et al., 1998) was higher than in the current study in all conditions (sometimes nearly three times as many stereotypes were used). However, since longer passages are likely to

contain more stereotypes, comparing these figures may not be all that meaningful. Macrae et al. (1994) used the stereotypic rating as the dependent variable, and the directed suppression condition essays were rated less stereotypic in all three studies compared with the control essays. The current results showed stereotypic ratings which were much lower for both placebo-directed suppression and placebo-control condition.

There are of course too many extraneous variables (such as passage length, the stimuli used, the instruction given to raters, and the stereotyped group) to draw definitive conclusions from comparing the current results to other studies. However it was suggested by Monteith at el (1998) that low-prejudice participants are less inclined to use stereotypes irrespective of task instruction. Therefore, the lack of statistical difference found in the current study with regards to suppression and the lack of stereotype use overall may be indicative of a lower prejudice population sample compared to those used in other studies, resulting in an overall floor effect. While strong social norms and adequate resources may also result in a minimal difference across conditions, the results of the HATH scale and the variation in positivity (discussed below) suggest that an overall low prejudice is the more likely cause. Of course, low prejudice and strong social norms against stereotyping are not mutually exclusive; in fact the occurrence of one is perhaps more likely to result in the other. The idea of an overall low prejudice will be explored further in the section on prejudice.

While no difference in overall stereotype use was observed in relation to directed suppression or the expected spontaneous suppression condition (glucose-control), these factors may have had an impact on the number of negative stereotypes used. Gailliot et al. (2009) found that among high-prejudice participants, those who

were given the glucose drink used fewer derogatory statements compared with the placebo condition. The current study employed a positivity rating instead. The positivity rating was expected to significantly correlate with this study's measure of prejudice (HATH score); however a significant correlation was not found. Although participants who were given the glucose drink had written more positive essays compared with those given the placebo drink (similar to the finding by Gailliot et al., 2009), this difference only occurred in the directed suppression conditions. Similarly, directed suppression resulted in higher positivity compared with no suppression instruction, but only in the glucose conditions, resulting in a significant interaction between the drink factor and suppression instruction. It is possible that participants interpreted the instruction to avoid using stereotypes as an instruction to avoid using negative stereotypes, which the glucose provided the energy to do so. However, since participants were not asked how they interpreted this instruction, it is not possible to conclusively determine this.

4.2 Post-Suppression Rebound

The interpretation and discussion of any post-suppression rebound effect in the current study is hampered by the fact that no overall suppression was observed. However, considering the low use of stereotypes overall, it is possible that the mental effort involved in suppression simply did not translate into a significant reduction in the number of stereotypes used in the writing task. Also, the significant interaction observed in the positivity rating may be indicative of suppression of negative stereotypes. Furthermore, according to Wegner's model of thought suppression it is the mental act of suppression that results in an increase in activation of the suppressed thoughts. Therefore, how successful that mental act is in reducing the observable use

of stereotypes may not matter with regards to the activation of those stereotypes. Despite no observed differences in stereotypicality in the writing task of the current study, the mental act of suppression could still have occurred and any post-suppression rebound effect should be explored.

As previously discussed, various studies (Galinsky & Moskowtiz, 2007, Macrae et al., 1994; Macrae et. al, 1998; Wyer et. al, 1998) have demonstrated that participants who are directed to avoid using stereotypes and participants who spontaneously suppress their use of stereotypes during an initial task tend to show a post-suppression rebound effect in subsequent tasks. Gailliot et al. (2009) demonstrated that an increase in glucose resulted in a spontaneous reduction in the use of stereotypes during a writing task. It was therefore expected that participants who were given a glucose drink would spontaneously suppress stereotypes during the writing task and would show a greater activation of gay stereotypes compared to non-suppressors. This increased activation of gay stereotypes was expected to manifest as a reduction in response time to identifying gay stereotypic words as real words in a lexical decision task compared with the response times of non-stereotypic (distracter) words. As such it was hypothesised that participants in the directed suppression conditions (glucose-directed and placebo-directed) and participants in the spontaneous suppression condition (glucose-control) would show a greater difference in response times to stereotypic and distracter words compared with participants in the non-suppression condition (placebo-control).

The current results did not support this hypothesis. As all participants in the main conditions were primed with an image and description of a gay male, it was expected that stereotypic words would be responded to faster than non-stereotypic words. The main effect of word-type did indeed show a significant difference with a

medium effect size. However, the between-subjects effects (beverage and suppression instruction), which were expected to cause variation in suppression of stereotypes and thus result in differences in post-suppression rebound effects, did not result in significant differences in response times. Furthermore, single factor analyses comparing all of the experimental conditions for each of the two word conditions showed there was no significant difference between conditions for either of the word types.

It is possible that due to the influence of social norms participants in the intended non-suppression condition (placebo-control) spontaneously suppressed their use of stereotypes resulting in a post-suppression rebound across all conditions. However, a certain degree of suppression in the placebo-control condition was expected, and the glucose should have better enabled participants to suppress stereotypes in the glucose conditions. Furthermore, the single factor analysis also included the baseline condition and showed that response times in this condition were not significantly different from any of the main conditions in either word type. As participants in the baseline condition were not primed with an image or description of a gay male, they were not expected to show an increase in activation of stereotypes. These results could be accounted for by participants in all conditions having been somehow primed with gay stereotypes outside of the experiment, therefore increasing stereotype activation in all participants. However, potentially more likely, is that participants in the experimental conditions had no increase in stereotype activation and therefore the difference in response times between word types was largely due to the specific words used for the LDT. The possible lack of activation is discussed further in the section on prejudice; whereas the effect of word type will be explored further in the limitations section.

The lack of any post-suppression rebound displayed in the suppression conditions is not consistent with other suppression studies such as Macrae et al. (1994), Macrae et al. (1998), and Wyer et al (1998), all of which found that suppressing thoughts in one task lead to a greater activation of those thoughts in subsequent tasks compared with people who did not suppress thoughts in the initial task. The rationale for this post-suppression rebound effect found in other studies came from Wegner's model of thought suppression, as described previously.

While the findings of the current study are not consistent with those from other post-suppression studies, it would not be reasonable to draw any implications from comparisons to the finding of the other studies or attempt to apply Wegner's model of thought suppression to the results found in this study, simply because suppression was not established in the initial task.

One factor which has been shown to contribute to both a low amount of stereotype use in all conditions as well as a lack of post-suppression rebound in stereotype activation, as has been seen in this study, is low prejudice.

4.3 Prejudice

Monteith at al. (1998) reported that participants who were instructed to suppress stereotypes in an initial task recalled more stereotypic words in a subsequent task. However, this only occurred with participants who were considered high in prejudice. Low-prejudice participants did not show a rebound effect. Similarly, Gailliot et al. (2009) found that giving participants a drink containing glucose reduced the number of derogatory statements made during a writing task about a gay male. Once again, however, this only occurred in high-prejudice participants as glucose had no effect on the already low number of derogatory statements made by low-prejudice participants.

These findings can potentially be explained by two separate theories. Low-prejudice individuals may not have stereotypic thoughts spontaneously activated following priming of a typically stereotyped target (Lepore & Brown, 1997). Alternatively, low-prejudice people may simply be more motivated and accustomed to suppressing stereotypes and thus have become more efficient at doing so without increasing the activation of stereotypes (Devine, 1989). Therefore, an overall low-prejudice sample would likely show little difference in suppression during the "day in the life" task and little activation of stereotypes during the LDT. This is what is believed to have occurred in the current study.

In the current study, explicit prejudice (as measured by the HATH scale) was not shown to correlate to stereotype use or positivity during the writing task, nor did it correlate with response times in the LDT. However, these analyses may be limited by the overall low-prejudice in all conditions resulting in a ceiling effect on the HATH scale (i.e., an overall low-prejudice floor effect among participants). The 20 item HATH scale has been widely utilised by a number of studies to measure prejudice (Gailliot et al., 2009; Gailliot, Plant, Butz, & Baumeister, 2007; Larsen et al., 1980; Monteith et al., 1998; Monteith, Deenan, &Tooman, 1996). The current study used a seven-point Likert scale, with a high score indicative of a more positive attitude towards homosexuals and homosexuality (low prejudice). The HATH score of participants in the current study did not significantly differ between conditions and an overall low mean score. Both Monteith et al. (1996) and Gailliot et al. (2009) used a seven-point Likert scale and had an equivalent mean score that was much lower, indicative of higher prejudice levels. This suggests a lower level of prejudice against homosexuals among participants in the current study. This lack of prejudice may come from various aspects of the sample population. It is possible that university

students, particularly those involved in social sciences (like Psychology) generally have more liberal views when it comes to homosexuals and homosexuality than the general population. It is also possible that the general New Zealand population have more liberal views compared with the general population of the United States, where most of the research in stereotyping has been conducted.

Furthermore, there is also evidence to suggest that cultural differences may have an influence on the occurrence of post-suppression rebound beyond the influence of prejudice. Zhang and Hunt (2008) reported that compared with U.S. participants, participants living in China showed a greater prejudice towards homosexuals but did not display a post-suppression rebound whereas American participants did. Similar results were obtained when the researchers used illegal immigrants as the target group. The researchers proposed the cultural differences beyond prejudice, such as reduced display of individuality and a higher motivation to obey authority figures may result in Chinese people having a more developed inhibition ability that leaves them less susceptible to post-suppression rebound effects. While New Zealand is generally considered more culturally similar to the U.S. than to China, especially with regards to displaying individuality, the Zhang and Hunt study demonstrates that post-suppression rebound effects may be culturally bounded.

4.4 Sex Differences

The sex of participants was another variable that was considered to have a likely influence on the dependent variables, particularly relating to stereotype use and prejudice. Larsen et al. (1980) reported that being male was a predictive factor of a low HATH score. In addition, LaMar and Kite (1998) found that while males held more negative attitudes towards homosexuals compared with females, they did not

differ with regards to endorsement of stereotypes. The results of the current study are consistent with the results of both of those studies.

Analysis of HATH scores revealed that sex was a significant factor with a medium effect size. Males scored lower on the HATH scale than females did, indicating that males held more negative views towards homosexuality. This is also supported by the fact that the "day in the life" passages written by males were rated as being significantly less positive than the passages written by females. Whereas the HATH scale indicates general attitudes towards homosexuals and homosexuality, the subject in the "day in the life" passage was a gay male, and therefore may reflect attitudes specific to homosexual males. LaMar and Kite (1998) found the heterosexual males held more negative attitudes about homosexual males than homosexual females, whereas heterosexual females held similar attitudes towards homosexual males and homosexual females except with regards to contact with homosexual people, where they rated potential contact with homosexual women more negatively than homosexual men. These findings suggests that if the subject in the "day in the life" passage had been a homosexual female instead of a homosexual male, the small effect size difference in positivity may have disappeared or been reversed. Consistent with LaMar and Kite (1998) males and females did not differ in the use of stereotypes during the writing task. Furthermore, the correlations between the dependent variables did not significantly differ between the sexes. These finding also illustrate the distinction between the endorsement of stereotypes and prejudice.

Sex differences were also observed with regards to the response times during the LDT. Females responded faster to all word types compared with males. There was no significant interaction effect between sex and word types, indicating that while females responded faster than males overall, the difference in response times between

stereotype words and distracter words were similar for both males and females. The overall faster response time of females during the LDT is consistent with other LDT studies (Weekes, Capetillo-Cunliffe, Rayman, Iacoboni, & Zaidel, 1999; Wegesin, 1998).

4.5 Limitations and Directions for Future Research

There are a number of limitations that need to be considered when interpreting the results of the current study. Attempts to overcome these limitations provide several possibilities that may be explored for future research. Firstly, a major limitation of this study was the low inter-rater reliability when examining the "day in the life" passages for stereotype use and positivity. All three of the measures of interest (proportion of stereotypes, stereotype rating, and positivity rating) had inter-rater reliability scores below .7. Previous research suggests that subjective measures such as stereotype rating and positivity rating are prone to low inter-rater reliability. Monteith et al. (1998) attempted to use a subjective measure of stereotypicality similar to that used by Macrae et al. (1994), but they reported that judges found the rating to be ambiguous and difficult and resulted in low inter-rater reliability. Monteith et al. (1998) instead opted for a more objective measure, looking at the proportion of stereotypic thoughts, and obtained a very high inter-rater reliability. In the current study a third rater was engaged to mark 40 of the passages, however inter-rater reliability across all variables remained low. While low-inter rater reliability on subjective measures can be understood, low reliability on the more objective measure (proportion of stereotypes) is harder to account for. While a list was provided of which exact stereotypes raters were suppose to look for, synonyms were also to be

included which may account for the discrepancy. Further training and practice for raters may be required in future to avoid low inter-rater reliability.

Another factor which may have influenced the lack of significant results was the specific words used in the LDT. It was expected that all participants who completed the "day in the life" passage would have gay stereotypes primed and would therefore be faster to respond to gay stereotype words than to the distracter words. However, this also occurred in the baseline group, which would suggest that the stereotype words were easier to identify as words compared with the distracter words. The study was limited by the number of participants that were available for recruitment, and running the baseline group first to test how easily identifiable the words are would expose the study to longitudinal variation. However in future it may be prudent to conduct tests to ensure stereotype words and distracter words are equally identifiable as words during an LDT prior to conducting the main study. This would likely be best achieved through matching the stereotype words to words with similar frequency of use.

While the HATH scale has been found to be a reliable measure of attitudes towards homosexuals (Larsen et al., 1980; Monteith et al., 1998), studies using this scale have mostly been conducted in the U.S. The overall low amount of stereotypes used and the high HATH scores and ceiling effect found in the current study may reflect an overall low prejudice towards homosexuals within the sample. As the sample was mostly comprised of university students studying stage one psychology, the sample may not reflect attitudes of the general New Zealand population Trialling the HATH scale on other population samples within New Zealand may indicate how appropriate it is for use with the New Zealand population. Furthermore, replicating this study on a sample with higher prejudice scores may increase the variation of

stereotype use and therefore allow for effects relating to directed suppression, spontaneous suppression, and post-suppression rebound to be better observed.

It could be suggested that the activation of social norms against stereotypes as well as instructions to avoid using stereotypes may have affected the participants' responses when completing the HATH scale at the end of the study. However, completing the HATH scale at the end of the experiment is consistent with other studies (Gailliot et al., 2009; Monteith et al., 1996; Monteith et al., 1998), the HATH scale is administered after the LDT and is therefore given in a context than the earlier "day in the life" passage, the HATH scale questionnaire encourages participants to be open and honest, and the HATH scale is considered to be fairly stable and not influenced by the activation of social norms (Monteith et al., 1996; Monteith et al., 1998).

It was intended that the two beverages provided would not differ in participants' estimation of calories or preference. While there was no significant difference in calorie estimation, the preference for the glucose drink over the placebo drink did approach significance. Studies have shown that people in a strong affective state (positive mood or negative mood) are more likely to use stereotypes when forming impression of other people compared with a neutral emotional state (Bodenhausen, Kramer, & Süsser, 1994; Esses & Zanna, 1995; Stroessner & MacKie, 1992). It is therefore possible that the better tasting drink would place participants in a happier mood that may result in an increase tendency to use stereotypes, which is supported by the significant positive correlation between participants rating of the drink and the stereotypic rating of the "day in the life" passages. At the same time, the increase in glucose could result in the person being better able to suppress stereotypes, and the two influences of positive mood and increased glucose may effectively cancel

each other out. However, while mood was not examined in the current study it has been measured in similar glucose studies and has not shown to be a significant factor on stereotype use (Gailliot et al, 2007; Gailliot et al., 2009).

Future studies should also consider the stimulus provided during the "day in the life" passage and the LDT. While using a homosexual male as the subject of the passages is consistent with other studies (such as Gailliot et al., 2009), the sex of the subject is a confounding factor when comparing male and female responses. Also, since a male was used in the "day in the life" passage, the stereotype words used in the LDT needed to be stereotypes specific to gay males. As discussed previously, heterosexual males respond more negatively to homosexual males than to homosexual females, whereas heterosexual females typically respond more positively and more similar towards homosexuals of either sex (LaMar & Kite, 1998). Therefore, considering there are four between-subject conditions in this study, it could be argued that there may not be a large enough sample size to identify interactions between suppression instruction and beverage sweetener, if these interactions are only occurring in participants of a specific sex. However, power analysis revealed this to be unlikely. The 2 (Sex) x 2 (Beverage) x (Suppression) ANOVA of the stereotype ratio had an observed power of .081 and a small effect size ($\eta_p^2 = .002$): therefore to obtain a significant interaction with 80% power, a sample size of at least 5400 would be required. An even larger sample size would be required to obtain a significant result with regards to the stereotype rating which had an observed power of .069 and an interaction effect size of $\eta_p^2 = .001$ The ANOVA examining the responses during the LDT with relation to sex, suppression instruction, beverage, and word type had an observed power of .051 and an effect size of essentially zero ($\eta_p^2 = .000$); therefore it would not be possible to determine a sample size that would be required to show a

significant effect. Therefore the lack of significance in the results is most likely due to the low effect size rather than an inadequate sample size. Future studies may find it useful to focus on recruiting only male participants when using a gay male target (as they are likely more varied in prejudice as well).

4.6 Implications and Applications

Due to the preliminary nature of these findings based on the New Zealand population, further research is necessary to be able to adequately draw comparisons of previous studies; therefore it is difficult to find many applications from the current results. However, the null findings do provide some implications for research into stereotype suppression particularly relating to boundaries of suppression and post-suppression rebound. While explicit prejudice showed no correlations to other dependent variables, the low level of prejudice overall and the lack of suppression, even when directed to suppress, supports the idea that stereotype suppression (whether spontaneous or directed) and subsequent rebound effects may require a comparatively higher level of prejudice to occur. That said, the fact that males and females did not differ in terms of stereotype use during the writing task or stereotype activation during the LDT, but did differ in positivity during the writing task as well as explicit prejudice (HATH), supports the notion that prejudice and stereotype activation are related but distinct aspects. These two finding may suggest that stereotyping does not increase prejudice, however prejudice may increase stereotyping. Interpreting any post-suppression rebound effects is primarily limited by the fact that no difference in suppression was observed in the initial task.

      As positivity in the writing task was shown to be higher when glucose was given in combination with the direction to avoid using stereotypes, this may have a

potential practical application. In workplace and school settings, for example, where it is not unheard of for notices to be placed and instructions given for attendees to practice tolerance, the findings of this study would suggest that the provided direction also requires attendees to have an adequately high blood glucose level for it to have an effect in reducing prejudice. Similarly, high glucose without direction does not seem to increase positivity. Although, these findings could be an artefact for overall low-prejudiced individuals, perhaps it would be more beneficial for such directional information and notices to be placed in areas where high blood glucose is likely to occur, as well as being more likely to be attended to (i.e. break rooms).

4.7 Conclusion

No significant results were obtained relating to the main hypotheses of this study. Overall suppression of stereotypes did not occur when participants were directed to avoid using stereotypes nor did it occur when participants were given glucose to aid in spontaneous suppression. However, higher positivity of the passages was found with participants who were both directed to avoid using stereotypes and were given glucose, suggesting that the combination of glucose and direction may assist in reducing the use of negative stereotypes. The experimental conditions did not differ with regards to the response times to stereotypic words and non-stereotypic words, and did not differ from the baseline condition. While prejudice was not found to correlate with any of the dependant variables, the overall low prejudice and low use of stereotypes is considered to be the main reason that no significant suppression of stereotypes or post-suppression effects were observed. Future studies should focus on recruiting participants with higher levels of prejudice to determine whether prejudice or other factors including cultural differences and mood limit the effects of stereotype

suppression in a New Zealand population. While significant findings similar to those found in other studies may have led this researcher to suggest maintaining sufficiently high glucose levels may benefit people by reducing stereotype usage and consequently improving interactions with target groups, the null findings do not exclude this as a possible benefit. However, future research may find that increasing blood glucose levels alone is of exclusive benefit to high-prejudice suppressors. Low-prejudice suppressors may not find the increased blood glucose reduces stereotype usage or prejudice unless also directed to avoid using stereotypes. The consequence that increased blood glucose may have on any post-suppression interactions has yet to be determined.

References

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The Influence of Afrocentric
        Facial Features in Criminal Sentencing. *Psychological Science, 15,* 674-679.

Bodenhausen, G. V., Kramer, G. P., & Süsser, K. (1994). Happiness and Stereotypic
        Thinking in Social Judgment. *Journal of Personality and Social Psychology,
        66,* 621-632.

Bodenhausen, G. V. & Lichtenstein, M. (1987). Social Stereotypes and Information-
        Processing Strategies: The Impact of Task Complexity. *Journal of Personality
        and Social Psychology, 52,* 871–880.

Castelli, L., Vanzetto, K., Sherman, S. J., & Arcuri, L. (2001). The explicit and
        implicit perception of in-group members who use stereotypes: Blatant
        rejection but subtle conformity. *Journal of Experimental Social Psychology,
        37,* 416–426.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and
        Conformity. *Annual Review of Psychology, 55,* 591–621.

Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression
        and suppression of prejudice: The struggle for internalization. *Journal of
        Personality and Social Psychology, 82,* 359–378.

Crocker, J., Voelkl, K., Testa, M., & Major, B. (1991). Social Stigma: The Affective
        Consequences of Attributional Ambiguity. *Journal of Personality and Social
        Psychology, 60*, 218-228.

Dalgleish, T., Yiend, J., Schweizer, S., & Dunn, B. D. (2009). Ironic Effects of
        Emotion Suppression When Recounting Distressing Memories. *Emotion, 9,*
        744-749.

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

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006).
        Looking Deathworthy: Perceived Stereotypicality of Black Defendants
        Predicts Capital-Sentencing Outcomes. *Psychological Science, 17*, 383-386.

Esses, V. M. & Zanna, M. P. (1995). Mood and the Expression of Ethnic Stereotypes.
        Journal of Personality and Social Psychology, 69, 1052-1068.

Fairclough, S. H., & Houston, K. (2004). A metabolic measure of mental effort.
        Biological Psychology, 66, 177–190.

Foschi, M., Lai, L., & Sigerson, K. (1994). Gender and Double Standards in the
        Assessment of Job Applicants. *Social Psychology Quarterly, 57,* 326-339

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D.
    M., et al. (2007). Self-Control relies on glucose as a limited energy source:
    Willpower is more than a metaphor. *Journal of Personality and Social
    Psychology, 92*, 325–336.

Gailliot, M. T., Peruche, B. M., Plant, E. A., & Baumeister, R. F. (2009). Stereotypes
    and prejudice in the blood: Sucrose drinks reduce prejudice and stereotyping.
    Journal of Experimental Social Psychology, 45, 288–290.

Gailliot, M. T., Plant, E. A., Butz, D. A., & Baumeister, R. F. (2007). Increasing Self-
    Regulatory Strength Can Reduce the Depleting Effect of Suppressing
    Stereotypes. *Personality and Social Psychology Bulletin, 33,* 281- 294.

Galinksy, A. D., & Moskowitz, G. B. (2007). Further ironies of suppression:
    Stereotype and counterstereotype accessibility. *Journal of Experimental Social
    Psychology, 43*, 833-841.

Kaiser, C. R. & Miller C. T. (2001). Stop Complaining! The Social Costs of Making
    Attributions to Discrimination. *Personality and Social Psychology Bulletin,
    27*, 254-263

Kelly, A. E., & Kahn, J. H. (1994). Effects of suppression of personal intrusive
    thoughts. *Journal of Personality and Social Psychology, 66,* 998–1006.

LaMar, L., & Kite, M. (1998). Sex Differences in Attitudes toward Gay Men and

Lesbians: A Multidimensional Perspective. *The Journal of Sex Research, 35,*

189-196.

Larsen, K. S., Reed, M., & Hoffman, S. (1980). Attitudes of heterosexuals toward

homosexuality: A Likert-type scale and construct validity. *Journal of Sex*

Research, 16, 245–257.

Lepore, L., & Brown, R. (1997). Category and Stereotype Activation: Is Prejudice

Inevitable? *Journal of Personality and Social Psychology, 72,* 275-287.

Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1998). Saying No to Unwanted

Thoughts: Self-Focus and the Regulation of Mental Life. *Journal of*

Personality and Social Psychology, 74, 578-589.

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.

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

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype

    threat: The effect of self-affirmation on women's intellectual performance.

    Journal of Experimental Social Psychology, 42, 236–243.

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.

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.

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-Control as Limited

    Resource: Regulatory Depletion Patterns. *Journal of Personality and Social

    Psychology, 74,* 774–789.

Robinson, S., Montiel, E., Jakubowski, B., & Madon, S. (1996). The content and

    strength of two stereotypes: Male republicans and male homosexuals. *Poster

    session presented at the annual meeting of the Eastern Psychological

    Association. Philadelphia, PA.* Retrieved April 16, 2009, from

    http://www.psychology.iastate.edu/faculty/madon/epa96hr.fin.htm.

Shapiro, J. R. & Neuberg, S. L. (2008). When Do the Stigmatized Stigmatize? The

    Ironic Effects of Being Accountable to (Perceived) Majority Group Prejudice-

    Expression Norms. *Journal of Personality and Social Psychology, 95,* 877 -

    898.

Shelton, J. N., Richeson, J. A., Salvatore, J., & Trawalter, S. (2005). Ironic Effects of

    Racial Bias During Interracial Interactions. *Psychological Science, 16*, 397-

    402.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype Susceptibility: Identity

    Salience and Shifts in Quantitative Performance. *Psychological Science, 10*,

    80-83.

Steele, C. M. & Aronson, J. (1995). Stereotype Threat and the Intellectual Test

    Performance of African Americans. *Journal of Personality and Social*

    Psychology, 69, 797-811.

Stoessner, S. J. & MacKie, D. M. (1992). The Impact of Induced Affect on the

    Perception of Variability in Social Groups. *Personality and Social Psychology*

    Bulletin, 18, 546-554.

Vorauer, J. D. & Turpie, C. A. (2004). Disruptive Effects of Vigilance on Dominant

    Group Members' Treatment of Outgroup Members: Choking Versus Shining

    Under Pressure. *Journal of Personality and Social Psychology, 87,* 384-399.

Weekes, N. Y., Capetillo-Cunliffe, L., Rayman, J., Iacoboni, M., & Zaidel, E. (1999). Individual Differences in the Hemispheric Specialization of Dual Route Variables. *Brain and Language, 67*, 110–133.

Wegesin, D. J. (1998). Event-Related Potentials in Homosexual and Heterosexual Men and Women: Sex-Dimorphic Patterns in Verbal Asymmetries and Mental Rotation. *Brain and Cognition, 36,* 73-92.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101*, 34–52.

Wegner, D. M., & Erber, R. (1992). The hyperaccessibility of suppressed thoughts. *Journal of Personality and Social Psychology, 63,* 903-912.

Wegner, D. M., Schneider, D. J., Carter III, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology, 53,* 5-13.

Wenzlaff, R.M., Wegner, D.M., & Rober, D.W. (1988). Depression and Mental Control: The Resurgence of Unwanted Negative Thoughts. *Journal of Personality and Social Psychology, 55,* 882-892.

Wyer, N. A., Sherman, J. W., & Stroessner, S. J. (1998). The spontaneous suppression of racial stereotypes. *Social Cognition, 16,* 340-352.

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.

Zhang, S. & Hunt, J. S. (2008). The stereotype rebound effect: Universal or culturally

bounded process? *Journal of Experimental Social Psychology, 44*, 489-500.

Appendix A

The Day in the Life with Directed Suppression

This task looks at how people form impressions of strangers based on limited information. You have been randomly assigned to one of 12 subjects, whom you will then be asked to write a brief passage about their daily life. Each subject has provided a photograph of themselves as well as a couple of sentences describing themselves.

"My name is Sammy. I'm 23 years old, live and work in Wellington. I'm gay and spend most of my free time hanging out with my mates, and I really enjoy going to the movies to relax".

On the paper provided, spend the next 5 minutes writing about a typical day in Sammy's life.

(Optional – Direct Suppression Instruction):

Previous psychological research has established that our impressions and evaluations of others are often biased by stereotypic preconceptions of groups that the person is a member of. In writing your "Day in the Life of…." passage we ask you to actively try to avoid using stereotypes of any groups that the subject in the photograph has identified him -or her-self as being a member of.

Appendix B

The 56 letter strings used in Lexical Decision task, the mean response time, and the

number of participants who incorrectly categorised each letter string (n =150)

Word Type	Letter String	Response Time (ms)	Errors
Stereotypes			
	understanding	787	3
	compassionate	837	9
	melodramatic	980	21
	affectionate	890	10
	fashionable	758	4
	hairdresser	779	10
	effeminate	1300	86
	flamboyant	839	24
	sensitive	688	8
	emotional	674	4
	talkative	760	7
	feminine	819	9
	artistic	750	8
	dainty	868	47
Distracters			
	uncharismatic	1268	31
	consequential	943	10
	invigorating	1007	30
	unproductive	938	6
	magnificent	767	4
	substantial	894	6
	compatible	914	12
	additional	708	3
	malleable	962	49
	efficient	772	9
	ingenious	938	21

freckled	880	26
terrific	710	6
clever	649	7
Non-words		
satisfunctary	1276	39
undercronking	1126	12
antiscolistic	1451	43
magnanimently	1324	21
unattrontive	1238	21
conquisitive	1296	78
unprolonging	1105	78
invenerating	1278	29
contricious	1179	38
constamable	1121	29
affactional	1149	34
submaritial	1352	62
effamulate	988	13
illubrious	1013	33
ingratious	1177	67
melocratic	1132	56
imperable	1061	74
farristic	953	9
tramitive	930	13
mendiable	1049	40
ellucient	1072	31
sensivate	1346	66
tossatic	840	10
femaline	1160	28
frookled	827	10
rompable	881	15
claber	857	12
bointy	826	19

Appendix C

Pilot study ratings of how much each trait was associated with heterosexual men and
homosexual men, as well as how positive the traits were. The difference refers to the
homosexual association less the heterosexual association. The p-value is the p value
of the t-test comparing the homosexual and heterosexual associations. The traits used
for the LDT were the traits that showed the greatest positive difference score, and
were statistically significant.

Trait	Homosexual Association		Heterosexual Association		Positivity Rating		Difference	p
	Mean	SD	Mean	SD	Mean	SD		
Effeminate	6.20	1.03	1.80	1.87	2.80	1.32	4.40	.000
Flamboyant	6.00	0.94	1.70	0.67	3.30	1.16	4.30	.000
Hairdresser	5.40	2.01	1.30	0.48	3.60	1.17	4.10	.000
Emotional	5.70	1.34	1.60	0.52	3.50	0.97	4.10	.000
Talkative	5.80	1.03	1.90	0.88	4.50	1.35	3.90	.000
Dainty	4.80	1.40	1.10	0.32	3.00	1.25	3.70	.000
Fashionable	5.70	1.16	2.10	1.10	4.80	0.79	3.60	.000
Melodramatic	5.30	1.77	1.70	0.48	1.70	0.67	3.60	.000
Feminine	4.80	1.48	1.30	0.67	2.70	1.16	3.50	.000
Artistic	5.40	1.07	2.10	0.88	5.10	0.88	3.30	.000
Sensitive	5.00	1.41	1.80	1.03	4.50	1.78	3.20	.000
Affectionate	5.50	1.08	2.40	1.07	5.80	1.14	3.10	.000
Understanding	4.90	1.37	1.90	0.74	5.90	0.74	3.00	.000
Compassionate	5.10	1.37	2.20	1.03	6.00	0.94	2.90	.000
Rainbows	4.10	2.16	1.20	0.32	3.20	1.14	2.90	.001
Neat	4.60	1.78	1.70	0.84	5.20	0.79	2.90	.001

Transvestite	4.20	1.99	1.40	0.52	2.50	1.43	2.80	.001
Tidy	4.50	1.84	1.70	0.70	5.30	0.95	2.80	.000
Sentimental	4.30	1.83	1.60	0.70	3.80	0.63	2.70	.000
Different	5.20	1.23	2.70	1.49	4.50	1.08	2.50	.000
Gentle	5.00	1.63	2.70	0.82	5.10	1.20	2.30	.000
Weak	3.80	1.93	1.60	0.84	2.40	0.97	2.20	.002
Individualistic	4.80	1.75	2.80	1.48	5.30	1.16	2.00	.000
Flashy	4.20	1.87	2.60	1.26	3.20	1.03	1.60	.015
Liberal	4.60	1.96	3.00	1.89	4.50	0.97	1.60	.026
Spunky	4.40	1.65	2.90	1.52	4.20	1.62	1.50	.005
Sociable	5.00	1.94	3.70	1.34	5.80	1.23	1.30	.002
Flirtatious	4.60	1.35	3.30	1.42	4.10	1.45	1.30	.005
Optimistic	4.30	1.77	3.10	1.37	5.80	1.14	1.20	.018
Enthusiastic	4.50	2.12	3.40	1.84	5.10	0.99	1.10	.077
Attractive	4.40	2.27	3.40	1.71	6.20	1.03	1.00	.031
Outgoing	4.70	1.42	3.70	1.34	5.20	0.63	1.00	.007
Thin	3.50	2.01	2.50	1.27	3.70	1.34	1.00	.021
Friendly	4.70	1.70	3.70	1.83	6.60	0.52	1.00	.037
Peculiar	3.10	1.79	2.10	1.73	3.40	1.65	1.00	.074
Intelligent	3.70	1.57	2.90	1.37	6.40	0.52	0.80	.026
Nice	4.20	1.93	3.40	1.35	5.70	1.34	0.80	.035
Manipulating	3.20	2.10	2.40	1.78	1.50	0.53	0.80	.197
Educated	3.70	1.42	3.10	1.52	6.60	0.52	0.60	.056
Wealthy	3.10	1.60	2.70	1.34	4.90	0.88	0.40	.187
Silly	3.10	2.02	2.80	1.87	3.60	0.84	0.30	.308
Stupid	2.60	1.58	2.40	1.71	2.00	1.15	0.20	.278

Hedonistic	4.30	1.57	4.20	1.81	3.60	1.07	0.10	.440
Outspoken	3.90	1.52	4.00	2.11	4.00	0.94	-0.10	.438
Unemployed	2.30	1.57	2.50	1.72	2.00	1.41	-0.20	.084
Cruel	2.10	1.20	2.40	1.35	1.00	0.00	-0.30	.234
Changeable	3.10	2.18	3.40	2.27	3.50	1.18	-0.30	.363
Ambitious	3.30	1.49	3.60	1.71	5.30	1.06	-0.30	.217
Perverted	3.40	1.65	3.80	2.20	1.80	0.92	-0.40	.267
Promiscuous	4.60	1.26	5.10	1.29	2.20	1.23	-0.50	.048
Shallow	3.70	1.95	4.20	2.35	1.90	1.29	-0.50	.136
Mean	2.60	1.65	3.10	1.79	1.60	0.70	-0.50	.149
Frank	4.50	1.58	5.00	1.41	4.80	0.92	-0.50	.181
Dodgy	3.30	2.00	3.80	1.81	1.90	0.88	-0.50	.232
Opinionated	4.30	1.57	5.10	1.85	3.90	1.37	-0.80	.076
Proud	3.70	1.77	4.50	2.12	4.30	0.95	-0.80	.124
Athletic	3.90	1.66	4.80	1.40	5.20	0.92	-0.90	.061
Fat	1.80	1.03	2.80	1.81	2.20	1.03	-1.00	.011
Unfriendly	2.00	1.56	3.00	1.49	1.50	0.71	-1.00	.026
Assertive	3.80	1.48	4.80	1.75	5.00	1.25	-1.00	.042
Serious	2.80	1.48	3.80	1.87	4.10	0.74	-1.00	.053
Informal	3.10	1.73	4.30	2.26	4.40	1.17	-1.20	.048
Intolerant	2.60	1.78	4.00	2.00	2.60	1.96	-1.40	.084
Stubborn	3.70	1.42	5.20	1.48	3.10	1.10	-1.50	.017
Prejudiced	3.00	1.25	4.60	1.35	1.40	0.52	-1.60	.002
Racist	2.10	1.10	3.80	1.55	1.30	0.48	-1.70	.001
Reckless	3.10	1.37	4.80	1.62	2.50	1.43	-1.70	.002
Conservative	1.80	0.92	3.70	1.34	3.60	0.84	-1.90	.001

Strong	3.00	0.67	5.10	0.99	5.50	1.08	-2.10	.000
Sloppy	1.90	0.74	4.30	1.49	1.90	0.74	-2.40	.000
Aggressive	2.50	1.18	5.00	1.41	2.00	1.15	-2.50	.002
Unshaven	2.00	1.05	4.50	1.96	3.10	1.45	-2.50	.003
Stoic	2.33	1.22	4.88	1.96	3.78	1.20	-2.54	.006
Traditional	2.10	1.45	4.80	1.99	3.70	1.06	-2.70	.001
Unemotional	1.80	1.03	4.70	2.00	3.40	1.43	-2.90	.004
Masculine	2.70	1.57	5.70	1.64	5.30	1.06	-3.00	.005
Rugged	2.00	0.82	5.30	1.42	5.00	1.41	-3.30	.000
Tough	1.80	0.63	5.20	1.14	4.70	1.06	-3.40	.000
Sexist	1.60	1.07	5.50	1.43	1.40	0.70	-3.90	.000
Macho	1.50	0.53	5.60	1.71	3.30	1.06	-4.10	.000

Appendix D

Sex differences on dependent variables

Measure	Males Mean (SD)	Females Mean (SD)	t	p	df	Effect size (d)
Beverage Rating	5.41 (0.90)	5.55 (1.00)	-0.797	0.427	118	.147
Calorie Estimation	97.52 (37.52)	104.33 (32.15)	-1.069	0.287	118	.195
HATH	5.61 (0.94)	6.12 (0.81)	-3.198	0.002	118	.581
No. of Stereotypes	0.975 (1.42)	0.958 (1.40)	-0.065	0.948	118	.012
Stereotype Ratio	.012 (.015)	.010 (.015)	0.621	0.536	118	.133
Stereotype Rating	3.52 (2.52)	2.97 (1.99)	1.328	0.187	118	.242
Positivity	5.88 (1.02)	6.39 (1.15)	-2.559	0.012	118	.469
Response Times (ms)						
Stereotypes	-.092 (.094)	-.145 (.107)	3.18	0.002	148	.526
Distracters	-0.077 (.113)	-.122 (.105)	2.565	0.011	148	.413
Non-Words	.005 (.128)	-.036 (.129)	1.956	0.052	148	.319