On-line elicitation of guilt using a prejudice paradigm

Nadine M. Kilchenmann
ACSENT Laboratory
Department of Psychology
University of Cape Town

Supervisor: Susan Malcolm-Smith
Co-supervisors: Kevin Thomas, Melike Nel

Word Count:
Abstract: [242]
Main body: [9994]

ABSTRACT

The current research serves as a pilot study to test the efficacy of a prejudice paradigm in the real-time elicitation of guilt, and to thus develop this paradigm as a reliable method for eliciting guilt in future neuroimaging studies. Previous research has shown that low-prejudice individuals report increased feelings of guilt when their *actual* reactions are in conflict with how they believe they *should* respond. Females between the ages of 18 and 30 years were recruited and pre-screened for suitability (e.g., they needed to meet specific criteria on measures of prejudice and social desirability, indicating that they have low prejudice and high concern for social desirability). Those selected were required to fill out further questionnaires and complete a set of Implicit Association Tests (IATs). Feedback following the critical set of IATs was pre-programmed and aimed at eliciting feelings of guilt by indicating to participants that they are prejudiced against certain racial and sexually-orientated groups. The results obtained from the experiment demonstrated that the manipulation of IAT feedback, with the purpose to elicit guilt, proved successful in 12 out of 19 cases. Further investigation of the data further showed that the effectiveness of the prejudice paradigm is highly dependent on a participant's personal characteristics and social environment. Thus, the current research strongly suggests that on-line elicitation of guilt, using the evaluated prejudice paradigm, can be achieved through careful participant selection.

Keywords: emotion elicitation; guilt; Implicit Association Test; moral emotions; prejudice; pride; shame.

Several psychological studies have defined a sub-category of emotions as the ‘moral emotions’. Such emotions differ from basic emotions in that they are thought to be the defining component involved in the apparatus that links moral standards to moral behaviour (Tangney, Stuewig, & Mashek, 2006). Within this theoretical framework, human moral behaviour is governed by a set of emotions, such as guilt and embarrassment, which guide our decision-making processes in terms of the initiation and inhibition of actions based on their moral acceptability. For instance, pro-social emotions such as guilt promote social conformity and reparative actions.

Tangney et al. (2006) consider moral emotions to be critically important in understanding an individual’s behavioural adherence or disobedience to socially accepted moral principles. The motivational force behind an individual’s conformity to social norms is, thus, based on moral emotions.

BACKGROUND

Whereas the basic emotions such as happiness, anger and fear have been the main focus of affective research, moral emotions have received little attention. One problem that makes studying the moral emotions more difficult is that they involve extensive cognitive processes based on the evaluation of the self (Lewis, 1992); in contrast, the study of basic emotions, such as fear, is less of a challenge as they can be elicited by specific, determinable stimuli. Although the elicitors of moral emotions can be broadly viewed as internal cognitive events, there is little agreement as to specifically what those elicitors might be. This lack of agreement largely arises because no specific stimuli have been found to invariably prompt feelings of guilt or shame, for instance. Because the immediate elicitor is cognitive in nature and no reliable cause-and-effect relationship exists between an external event and the experience of a moral emotion, it remains a challenge to consistently and unfailingly elicit such emotions in a controlled setting (Lewis, 2000).

The little research that has been conducted on moral emotions has primarily focused on the negatively associated moral emotions of guilt, shame and embarrassment as well as the positive feeling of pride. Although these emotions vary according to the emotion-eliciting situation, they are similar in that elicitation occurs as a reaction to some behaviour that morally evaluates the self (Tangney et al., 2006).

Theories of emotion

A variety of theories attempt to explain the experience of emotion. One of the earliest models of emotion, known as the James-Lange theory, emphasises the physiological determinants of emotion. For instance, Lange (1885) suggested that the conscious experience of emotions results from one's perception of different autonomic activations. In contrast, the Cannon-Bard model argued that physiological arousal can occur in the absence of an emotional response (Cannon, 1927). Schachter's (1964) two-factor theory, however, proposed that emotion is deduced from an autonomic arousal and is labelled as a certain emotion following the cognitive explanation of the stimulation. The above theories oppose the evolutionary theory of emotion, which describes emotions as innate reactions requiring little cognitive evaluation. For instance, Panksepp (2004) suggests that "emotional feeling may reflect the neurodynamics of brain systems that generate instinctual emotional behaviours" (p. 30).

Although there is a wealth of contemporary psychological research on emotion, researchers have not conclusively settled on a common theoretical framework; debate surrounding the validity of each of the abovementioned theories is, to a large degree, ongoing. Generally, however, researchers characterize emotions as complex experiences, arising from cerebral processes, that can be divided into different dimensions depending on how they are expressed and elicited. In this regard, affective neuroscience has attempted to divide emotions into different categories, namely primary and secondary emotions, based on the different brain regions involved in their elicitation.

Primary emotions, such as fear, disgust and happiness are regarded as innate responses to sensory experiences. These prototypical emotions can be triggered by specific, determinable external events and are recognizable with the help of universal facial expressions. Furthermore, they are organised in a hierarchical structure in executive operating systems in the brain that act with distinct neural circuits and neurochemicals. Primary emotions appear to govern higher-level brain functioning such as decision-making processes (Damasio, 1995) and thus also influence secondary emotions.

The more complex secondary emotions, such as pride, guilt and embarrassment, are acquired through learning and experience and are uniquely human (Panksepp, 1998). These social, self-conscious emotions are extensively internalized and are triggered as a response to cognitive self-evaluations; they are not associated with any characteristic facial expression,

unless co-occurring with a more basic emotion. An example of the latter would be the pairing of the internalized feeling of pride with a facial expression relating to joy, such as a smile. Although pride may be related to joy it is the latter that triggers the facial cue. Thus, in this case, pride could only be assessed via a self-report measure.

Moral emotions

Moral or social emotions, such as shame, guilt, embarrassment and pride can be further grouped into a category known as the self-conscious emotions. The characteristic feature of this group is that they are evoked by self-reflection and self-evaluation. Such self-evaluation occurs as implicit or explicit conscious experiences, and involves a subconscious process integrating the self. As self-reflection occurs, this group of emotions offer immediate feedback on behaviour and may take on a punishing or reinforcing role in terms of the event's moral acceptability (Tangney et al., 2006).

In contrast to lengthy deductive reasoning, self-conscious emotions therefore offer a rapid and unconscious evaluation of possible consequences relating to certain behaviours (Haidt, 2001). For example, feelings of shame may indicate that a specific action was wrong based on the individual's moral evaluation of the situation. As a result, one can attempt to rectify the moral wrong-doing by the initiation of a 'counter behaviour' such as an apology. In other words, these emotions help in upholding social standards in that they encourage reparative behaviour such as confession (Berthoz, Gre`zes, Armony, Passingham, & Dolan, 2006). In this regard, feelings of guilt, for example, are elicited by negative associations in response to behaviour deviating from morality (e.g., behaviour that harmfully affects other individuals), whereas positive feelings of pride and self-approval result from morally acceptable behaviour that leads to a positive evaluation of the self (Tangney et al, 2006).

The moral emotion of guilt

So far the discussion has established that several emotions, including guilt, shame, and embarrassment, are assumed to play a fundamental role in social morality. Although research on guilt is considered important in psychoanalytic theory as well as in the domain of socialization (see, e.g., Amodio, Devine & Harmon-Jones, 2007), in the 1970's and 1980's it was virtually

ignored by social and developmental psychologists. However, since the 1990's research has increasingly begun to focus on self-conscious emotions (Eisenberg, 2000).

In contemporary conceptions, the type of guilt relevant to moral behaviour is guilt that is a reaction to regretting a wrongdoing (Eisenberg, 2000). Within this context guilt may be defined as ''an agitation-based emotion or painful feeling of regret that is aroused when the actor actually causes, anticipates causing, or is associated with an aversive event" (Ferguson & Stegge, 1998, p.667). In this regard the guilty actor accepts responsibility for behaviours violating internalized standards of the self or others and then attempts to punish him/herself.

Differentiating moral emotions

The vast majority of research concerning moral emotions has focused on two negatively valenced self-conscious emotions, namely guilt and shame. In the past, the moral emotions of shame and guilt have been used synonymously because they share some common characteristics. For instance, they both involve an individual recognizing that s/he has violated a moral standard. Additionally, guilt and shame are equally likely to occur in response to the same situation: For example, one may feel guilty for having lied as well as feeling shamed if caught doing so and being confronted about it.

Recent attempts have been made to differentiate the two emotions by depicting them as two separate experiences. Distinguishing between guilt and shame can be arranged into three different categories: a distinction based on (a) the type of the eliciting event, (b) the public versus private nature of the transgression, and (c) the extent to which the person interprets the emotion-eliciting event as a failure of behavior or of the self (Tangney et al., 2006).

Empirical research has shown, however, that the type of eliciting event is of little relevance when it comes to distinguishing shame and guilt. Analyses of experiences of shame and guilt, as described by both children and adults, exposed no 'typical', distinctly unique shame or guilt-inducing situations (Keltner & Buswell 1996; Niedenthal, Tangney, & Gavanski, 1994; Tangney, Wagner, & Gramzow, 1992; Tracy & Robins, 2006). Some people relate events such as lying, cheating and stealing to shame, while others associate the same situations with feelings of guilt. However, some studies have shown that shame is evoked by a more extensive range of events (including both moral and non-moral transgressions, such as socially inappropriate behaviour or failure in performance situations), whereas guilt appears to occur exclusively in

relation to moral transgression (Eisenberg, 2000; Ferguson, Stegge, & Damhuis, 1991; Smith, Webster, Parrot, & Eyre, 2002).

A further important difference between guilt and shame is alleged to be the public versus private context of transgression (e.g., Benedict, 1946). Shame is regarded as the more "public" emotion, a response to public exposure and disapproval of some weakness. Guilt, on the other hand, is regarded as the more "private" experience, which arises from self-generated dilemmas of the conscience. However, empirical research has not yet been able to support this public/private distinction in terms of the nature of the emotion-eliciting event: Analyses of the social context of shame and guilt-eliciting situations has illustrated that both emotions are equally likely to occur in the presence of others (Niedenthal et al., 1994; Tangney, Miller, Flicker, & Hill- Barlow, 1996).

Another frequently cited distinction between shame and guilt is the degree of focus on the entire self versus a person's behaviour. Shame manifests itself in a negative evaluation of the entire self where the inappropriate behaviour, the 'bad' thing, is experienced as reflection of the "bad self" and the entire self is negatively evaluated. Thus, during the experience of shame the entire self is exposed and becomes inferior. During feelings of guilt, however, a specific behaviour or failure to act is negatively evaluated and remorse is only felt over the "bad behaviour"; the entire self is not scrutinized (Lewis, 1971; Tangney, Wagner & Gramzow, 1992). Consequently, guilt is reported to be a less painful emotional experience than shame because during feelings of shamefulness one's core self, and not simply one's behavior, is at stake (Tangney, Niedenthal, Covert, & Hill-Barlow, 1998). Guilt and shame can, therefore, be distinguished in that guilt is linked to feelings of remorse and regret but does not affect an individual's core identity. Therefore, whereas shame is associated with the urge to undo components of the self, guilt is related to the desire to undo a behavioural aspect (Niedenthal et al., 1994).Thus, shame rather than guilt is more likely to result from morally inappropriate situations involving concern about others' evaluations of the self, while situations that may be followed by reparative behaviour may be determinants of guilt (Ferguson et al., 1991; Tangney et al., 1992). These two moral emotions can, therefore, be differentiated in terms of "egocentric" versus "other-oriented" concerns (Niedenthal et al., 1994).

The self-conscious emotion of embarrassment, however, appears to be less central to the domain of morality than shame and guilt. For instance, during events that elicit embarrassment

individuals are less concerned with the morality of the situation than when feeling shame and guilt (Tangney et al., 1996). Nonetheless, in some situations embarrassment may influence people's conformity to moral standards (Tangney, Stuewig & Mashek, 2006).

Whereas pride is also categorized as a self-conscious emotion, unlike shame, guilt and embarrassment, it is regarded as a positively valenced moral emotion. Although pride involves self-evaluation the appraisal of the self is related to a socially valued outcome. As pride seems to enhance people's self-worth this positive social emotion may promote behaviour that meets social standards. For instance, feelings of pride for to conforming or exceeding important moral standards may serve a motivational purpose, reinforcing moral behaviour and inhibiting immoral impulses (Tangney, Stuewig & Mashek, 2006).

Although shame, guilt, embarrassment and pride vary in valence they, nonetheless, remain grouped within the same category of self-conscious emotions. This categorization is valid as the emotions are similar in that in each case the emotion is elicited when some aspect of the self is examined and assessed with regards to moral standards.

Elicitation of self-conscious emotions in the laboratory

Although moral emotions have been the topic of recent studies, reliably eliciting a particular moral emotion such as guilt, in experimental settings, remains a difficult proposition. One great challenge in research involving elicitation of moral emotions is the subtleness of facial expression during the experience of guilt and other social emotions, which makes their detection extremely difficult. According to Ekman (1999) because an individual can inhibit the external appearance of an emotion, emotions can occur without any facial signal. He further argues that a certain threshold may need to be crossed in order to bring about a visible emotional reaction. Therefore, the detection of moral emotions often relies extensively on subjective evaluation such as self-report, which is not as reliable as scientific measurement. Therefore, in order to be able to research moral emotions a reliable technique needs to be developed that is capable of consistently eliciting the emotion of interest.

Another complicating factor specifically in guilt elicitation is that guilt and shame often co-occur. In moral transgression, for example, a person may feel guilty for disobeying a social standard, while at the same time feeling shameful about his/her own shortcomings (Takahashi, 2004).

The above issues are of concern because without a way to reliably and distinctly elicit the different moral emotions there is no way to reliably study them in a controlled laboratory setting. Previous studies have used unsatisfactory means to elicit moral emotions such as guilt. For instance, Shin et al. (2000) used script-driven imagery in their study to elicit the emotional experience of guilt. In this paradigm participants were presented with recorded descriptions of personal events that they had written down previously. One of the described emotional events had to be one where the participant had felt the most guilt s/he ever experienced. Participants were then required to recall those events and imagine the related emotions they felt at the time. It appears controversial to accept the act of remembering feelings of guilt to represent the same emotional experience as actively experiencing the emotion.

In another attempt to elicit moral emotions in a laboratory setting, Britton et al. (2006) hypothesized that this process may be dependent on the sociality of the stimulating context. They exclusively used film and picture stimuli, varying in sociality (non-social versus social) and valence (negative versus positive emotions), as the emotion-eliciting mechanisms. A major problem with this approach, however, is that an emotional response to a passive stimulus (such as a visual image) is likely to be considerably different in emotional intensity compared to an emotion being currently experienced as a result of active participation in an emotion-eliciting situation.

A further example of an elicitation controversy is revealed by the study of Berthoz et al. (2006). Participants were required to rate socially inappropriate behaviour described in stories. Four types of verbal material were used where the protagonist of the story was either the participant or someone else and the violation of social norms was described as being either accidental or deliberate. Although these accounts included the individual themselves through personal reference, no personal active moral violation took place at the time. Rating oneself based on fictitious representations is likely to yield a different evaluation compared to ratings based on actual personal current experience.

The common underlying problem with the elicitation studies described above is that none of them elicited a moral emotion on-line. It seems that the key to true, purposeful, and intense moral emotion elicitation is to create an ecologically valid situation in which the participant actively experiences the required emotion as a real feeling during that very moment.

Specific Aims

The current research involves the administration of a prejudice paradigm in the elicitation of emotions especially that of guilt. Implicit Association Tests (IATs) presented within such a prejudice paradigm are the main instrument attempting to actively elicit the moral emotion of guilt. Taking into account the highly social and moral nature of the paradigm the elicitation of other moral emotions is not ruled out. It is hypothesized that under the application of the prejudice-negative IAT component a cluster of self-conscious emotions, namely guilt, shame and embarrassment will primarily be elicited. Furthermore, it is expected that the prejudice-positive IAT component will primarily elicit positively valenced moral emotions such as pride and satisfaction.

The research question under investigation may thus be stated in this way: How effective is the systematic manipulation of IAT feedback, within a prejudice paradigm, at reliably eliciting guilt and other moral emotions on-line?

METHOD

Design and Setting

The current research is a within-subjects social psychological experiment. Participants, recruited specifically because they had explicitly proclaimed they were not prejudiced and had shown themselves to be in need of the social approval of others, were deceived by bogus IAT feedback into believing that they were prejudiced toward individuals of particular races and sexual orientations. Data were collected through questionnaires administered both before and after the IATs. The main experimental procedures took place in the ACSENT laboratory in the Department of Psychology at the University of Cape Town.

Participants

A large group of undergraduate students (N = 147) completed a web-based survey that acted as a screening measure. Based on the results of this screening survey and the satisfaction of certain demographic inclusion criteria, 19 individuals were selected to participate in the main experimental procedures. The inclusion criteria specified that the participant be female, white,

not Jewish[1], and between the ages of 18 and 30 years. A female-only sample was used to reduce possible gender-related variability in physiological responses (Shields, 1991). The other demographic inclusion criteria were established in order to ensure a homogenous sample in terms of race, religion, and age (all factors important to the success of the experimental manipulation).

All participants (including those who only completed the web-based survey) received course credit via the Department of Psychology's Student Research Participation Programme. All study procedures were approved by the Research Ethics Committee of the UCT Department of Psychology. All participants gave informed consent and confidentiality was assured prior to the experiment.

Measures

Screening measures

Three screening instruments, administered through a web-based survey, aimed at identifying those participants who (a) held low-prejudice attitudes towards certain social groups, and (b) placed high value appearing socially appropriate in that regard. All of these instruments (the Marlowe-Crowne Social Desirability Scale (MCSDS; Crowne & Marlowe, 1960), the Internal Motivations Scale/External Motivations Scale (IMS/EMS; Plant & Devine, 1998) scale and the rating thermometer) have been frequently used in social psychological studies over the past several decades (e.g., Glaser & Knowles, 2008; Hawkins, 1968; Herek, 2000).

The MCSDS is a 33-item scale that measures the tendency of an individual to present an unrealistically positive impression of themselves. It thus assesses the extent to which someone is willing to give what that individual deems to be a socially appropriate response rather than an honest response. The MCSDS has been used in numerous studies as an indicator of social desirability (e.g., Carstensen & Cone, 1983; Kozma & Stones, 1987). Millham (1974) hypothesized that females who score high on the MCSDS will attempt to avoid social disapproval by avoiding behaviours that may imply that they have undesirable personal qualities. Thus, a female sample scoring high on this scale seems well suited for the current research. The

[1] One of the critical IATs investigated prejudice against religious groups, with Judaism being one of them.

set of items appearing on the full 33-item MCSDS has good internal consistency (Cronbach's alpha = 0.73; Barger, 2002).

The IMS/EMS is a 10-item rating scale concerned with various personal and social reasons people may have for trying to respond in a non-prejudiced manner towards different social groups. Participants indicate their level of agreement with each item on a scale ranging from 1 *(strongly disagree)* to 9 *(strongly agree).* Items are reverse coded when necessary so that high scores on each scale reflect higher levels of that type of motivation. The IMS/EMS shows good convergent and discriminant validity with reasonable alpha levels, ranging from 0.76 to 0.85 (Plant & Devine, 1998). For the current research, the rating questionnaire was aimed at determining attitudes towards four different groups (black, fat, Jewish and homosexual individuals). These groups were used in this study's IAT protocol.

The rating thermometer (see, e.g., Herek, 2000) required participants to give a single attitude rating, between 0˚ and 100˚, towards each of the four different groups described above. They were instructed that a thermometer rating of 0˚ would indicate an extremely unfavorable attitude toward a particular group, whereas a rating of 100˚ would indicate an extremely favorable attitude towards that group.

Experimental measures

Selected participants were initially asked to complete a set of questionnaires. All of these questionnaires, apart from the biographical and general health questionnaire, which was developed specifically for use in this study, are commonly used in psychological studies (see, e.g., Crawford & Henry, 2004; Johnson, Turner & Iwata, 2003; Knecht et al., 2000; Moilanen, 1993; Rusch et al., 2006; Stones & Kozma, 1994). A brief description of each of the measures is given below.

The Positive and Negative Affective Schedule (PANAS; Watson, Clark & Tellegen 1988) is a 20-item scale with 10-item positive affect and negative affect subscales that measure general mood/affect. Participants respond on a 5-point Likert-type scale indicating the extent to which they normally feel a certain mood/affect, such as being excited or upset. The PANAS scale generates two relatively independent affect dimensions: a positive and negative affect score. The overall likelihood of reporting somatic sensations associated with any emotion are usually higher in individuals with higher negative affect scores (Coan, Allen, & McKnight,

2006). The PANAS has been validated in several settings and has generally been shown to be reliable and consistently reflective of the correlating dimensions of PA and NA (DePaoli & Sweeney, 2000; Melvin & Molloy, 2000). Cronbach's alpha for the 10-item PA subscale and 10-item NA subscale has been shown to be 0.85 and 0.82, respectively (Thompson, 2007).

The Affect Intensity Measure (AIM; Larsen, 1984) is a 40-item scale designed to measure the intensity with which individuals typically experience positive and negative emotions (Larsen, 1984). Participants respond on a 6-point Likert-type scale agreeing or disagreeing with statements such as "When I receive an award I become overjoyed", "I feel pretty bad when I lie", "When I do feel anxiety it is normally very strong," and "Sad movies deeply touch me." The utility of this instrument is largely based on the fact that wide individual differences appear to exist in the characteristic intensity of affective response to the same emotion-evoking event. Evidence has attested to both the reliability and validity of this instrument, with documented Cronbach's alphas in the range of 0.90 to 0.94 (Larsen & Diener, 1987).

The Behavioural Inhibition System/ Behavioural Approach System (BIS/BAS; Carver & White, 1994) is a 20-item scale measuring behavioural inhibition and behavioural activation. A behavioural approach system (BAS) is said to regulate appetitive motives, where the goal is to move toward something desired. The behavioural inhibition system (BIS) is believed to regulate aversive motives, where the goal is to move away from something unpleasant (Carver & White, 1994). Participants respond to each of 22 statements on a 4-point Likert-type scale ranging from 1 (*quite untrue of you*) to 4 (*quite true of you*). The BIS/BAS possesses good reliability and Cronbach's alpha for the BIS and BAS subscales has been shown to be 0.76 and 0.83, respectively (Jorm et al., 1999).

The Beck Depression Inventory – Second Edition (BDI-II; Beck, 1996) is a 21-question multiple-choice self-report inventory measuring the severity of depression. Its items relate to depression symptoms such as hopelessness and irritability as well as cognitions such as guilt or feelings of being punished. Physical symptoms such as fatigue, weight loss, and lack of interest in sex are also considered (Moilanen, 1993). The BDI-II exhibits a high level of internal consistency, with alphas ranging from 0.89 to 0.93 (Whisman, Perez, & Ramel, 2000).

Implicit Association Test

Research in social cognition has suggested that many cognitive processes that influence human behaviour are not consciously controlled. The IAT (Greenwald, McGhee, & Schwartz, 1998) is a commonly-used social psychological experimental paradigm that measures these implicit influences on behaviour by assessing the strength of automatic associations between representations of mental concepts. Associations involving two pairs of contrasted concepts are measured indirectly and are therefore difficult to fake (Greenwald et al., 2002).

The IAT has been used to measure attitudes towards various environmental objects, as well as self-esteem and prejudice. Various researchers agree that implicit cognitive measures can be differentiated from self-report in that they expose mental associations without having introspection as a requirement (Greenwald & Banaji, 1995; Wilson, Lindsey, & Schooler, 2000). A recent meta-analysis by Nosek, Greenwald, and Banaji (2005) showed that the IAT is a better predictor of some types of human actions than more conventional overt self-report methods. Additionally, the IAT has contributed to the study of implicit social cognition due to its ease of administration, large effect sizes and good reliability (Nosek et al., 2005).

The prejudice paradigm administered in the current research relies on a central finding from social psychological research into prejudiced attitudes: frequently, people who claim to hold non-prejudiced attitudes show prejudiced tendencies on non-consciously monitored measures such as the IAT (Crosby, Bromley, & Saxe, 1980; Gaertner & Dovidio, 1986). Primarily, these contradictory responses appear to emerge from the fact that conscious decisions to reject prejudice do not ultimately result in the complete elimination of prejudiced responses (Devine, Monteith, Zuwerink & Elliot, 1991). Thus, overt and explicit non-prejudiced beliefs and attitudes may, within the same individual, coexist with covert and implicit prejudiced beliefs and attitudes.

In the current study, participants were presented with false feedback to some of their IAT responses. This false feedback indicated that, despite their explicit self-reports, they are indeed implicitly prejudiced towards certain social groups. Given the research by Devine et al. (1991), the possibility of such a situation occurring within the research context seems plausible.

Furthermore, Britton et al. (2006) emphasize that the emotion-triggering process may be dependent on the sociality of the stimulating context. The context of the current study was highly public and social, in that the participants were informed that their IAT results were visible to, and

recorded by, the researchers. According to Frijda (1994) guilt is "usually associated with the belief that one has harmed another person" (p. 43). Thus, we anticipated that participants would feel that, because of their IAT results, they had done something morally wrong because they were showing prejudice. Therefore, it was expected that participants would experience the self-conscious emotion of guilt as the IAT feedback led them to believe that they had committed a moral violation.

The current study required devising a specific IAT protocol for six different social groups all of which were tailored with pre-programmed response feedback to suit the projects' purposes. More specifically, there were a total of 6 IATs divided into three different categories (neutral, prejudice-positive, and prejudice-negative). Within each of the three sets of IATs there were two topics, each of which was presented only once within the entire session. The topics for the neutral IATs were sports (preference for *swimmers* versus *runners*) and facial hair (preference for *facial* versus *no facial hair*). The topics for the prejudice-positive IATs were weight (preference for *fat* versus *thin* people) and religion (preference for *Judaism* versus *other religions*). The topics for the prejudice-negative IATs were sexuality (preference for *gay* versus *straight* people) and race (preference for *black* versus *white* people).

The feedback for the neutral IAT, unlike that of the prejudice-positive and prejudice-negative IATs, was not manipulated to give pre-programmed feedback. Instead the results of the neutral IAT were real indicating the participant's true implicit attitude towards these groups. . The neutral IATs were not expected to elicit any strong positive or negative affect, since these categories were not controversial in a social sense in hence any preference would only be a matter of personal choice.

The IATs crucial to the elicitation of self-conscious emotions were the two suggesting negative-prejudice attitudes. All participants received the same manipulated feedback which falsely led them to believe that the answers they gave were highly prejudiced (see Appendix A). This harshly suggested that the participant severely discriminates against black and homosexual people although the pre-screening had shown that she is not prejudiced against these groups. The combination of the negative feedback and the awareness that the researcher sees the prejudiced results immediately was anticipated to be effective in eliciting feelings of guilt in the participants.

In the prejudice-positive IATs the feedback was pre-programmed to be positive i.e. indicating no prejudice tendencies towards the specific groups. It was argued that receiving feedback indicating that you respect people from various religions equally (religion IAT), and likewise show no prejudice against fat people (weight IAT) would induce positive emotions of pride and satisfaction in participants with high social desirability. These IATs were therefore critical for the elicitation of positive moral emotions. At the same time these IATs also served as a precautionary measure preventing the participants from questioning the validity of the IATs as it might appear suspicious if prejudice-confirming feedback was presented after every response.

Emotion checklist

The main experimental control measure comprised a self-report emotion checklist created specifically for this study (see Appendix B). The checklist, which was administered after each IAT, recorded the emotional response of the participants to the elaborated feedback presented following each IAT. In other words, we used this checklist to measure which moral emotions were elicited by the IAT feedback.

Statistical analysis was done using Statistica version 8 software package (StatSoft Inc., 2007).

Procedure

The study required some preliminary preparation and distinct selection of participants from a larger sample and was therefore divided into three phases: (1) the pre-screening and selection of participants, (2) contacting the selected participants, (3) the actual laboratory experiment.

Phase 1: The pre-selection process

The eligibility of survey participants to continue on to subsequent phases of the experiment was contingent upon them fulfilling certain inclusion criteria. The primary criterion that needed to be fulfilled was a thermometer rating of 60 and above on the critical categories (*black* and *homosexual*), as well as high thermometer ratings for the remaining two categories (*fat* and *Jewish*).

A second inclusion criterion was that participants needed to score above the sample median on the MCSDS. Those who scored slightly below the median but met the other two inclusion criteria remained in consideration, however.

The third and final inclusion criterion was based on an evaluation of individual IMS/EMS results. For this purpose, the IMS scores were ignored altogether and the focus was on the EMS scores, which indicate external/social motivations for non-prejudiced responses. To be eligible for further participation, individuals ideally had to score above the sample median for each category. Those who scored below the sample median on the critical categories ('black' and 'homosexual') were immediately excluded. Those who scored below the sample median on the non-critical categories, but above the sample median on the critical categories, remained in consideration for further participation if they met the other two inclusion criteria.

In summary, we chose these inclusion criteria because individuals who met them would have (a) self-reported (on the rating thermometer and EMS) favorable attitudes towards the relevant social categories, suggesting low prejudiced attitudes, and (b) self-reported (on the MCSDS) a high level of socially desirable responses, indicating a need for social approval from others.

The responses of each web survey participant were assessed in order to determine the degree to which the inclusion criteria were met. Those survey participants whose scores showed the greatest overlap with the inclusion criteria across all three questionnaires were chosen for the follow-up study. In this way, 19 individuals were identified as being suitable for further participation in the study.

Phase 2: Telephonic screening

The 19 individuals eligible for further participation were contacted telephonically. The phone call aimed to establishing that the selected participants were white females who were not on any psychoactive medication, had no history of a psychiatric disorder or psychiatric treatment, and had no history of head injuries or neurological disorders. Candidates who did not meet these criteria were not eligible for further participation in the study. Participants who met all the criteria were asked to indicate their willingness to take further part in the research; if they indicated their agreement, a suitable time slot was arranged for Phase 3 of the study.

Phase 3: Laboratory experiment

Each participant was individually assessed in a psychology laboratory. Upon arrival at the laboratory, the participant was asked to read and sign an informed consent document. Participants were informed that the purpose of the study was the investigation of prejudiced attitudes among UCT students. The true nature of the experiment was concealed because knowing that the aim of the study was to elicit guilt might have invalidated the participants' responses due to demand characteristics (see Appendix C for cover story). The actual intention of the study was only revealed at the end of the assessment, during a thorough debriefing.

After completion of the informed consent (see Appendix C) participants were required to complete the set of questionnaires described above (AIM, BDI-II, BIS/BAS & PANAS). After the participant had completed those questionnaires, the experimenter introduced her to the IAT section of the protocol.

The IAT task was explained to the participant with the help of an informative and detailed slide presentation on a computer monitor. Each slide of the presentation was timed to ensure that each participant received the same amount of exposure to the study material. The presentation also aimed at introducing the different social concepts the participants would encounter during the IATs. Prior to starting the study IATs, participants were required to do a practice IAT, on the Project Implicit website, to familiarise them with how the task worked. They were encouraged to approach the experimenter if anything was still unclear. Crucially, participants were informed that the researcher would receive immediate feedback, via a monitor on the other side of the laboratory, on their IAT results.

The IATs were pre-programmed in a pseudo-random order and counter-balanced across all participants. Five IAT sequences were created so that every 6th participant received the same combination of IATs. The combination of IATs differed from one participant to the next in terms of the order in which each topic of each IAT category (i.e., neutral, prejudice-positive or prejudice-negative) was presented (see Figure 1). Each sequence always started with a neutral IAT (sports IAT) followed by a random order of IATs. For instance, if participant 1 received the IATs for sports and weight after each other, then for participant 2 the sequence changed to sports followed by sexuality.

A C F B D E,

A F B D E C,

A B D E C F,

A D E C F B,

A E C F B D.

Figure 1. Pseudo-random IAT sequence: Each letter represents a different IAT.

After each IAT the participant received a brief 10 second feedback slide on her performance. This feedback was followed by a slide, presented for 20 seconds, giving more elaborate and detailed feedback about the tendencies and attitudes revealed by the previous IAT. This elaboration and detail aimed at strengthening any particular emotional experience felt in response to the brief feedback (see Appendix A).

After the 20-second elaborated feedback period, the participant was required to complete a self-report emotion checklist (see Appendix B) to establish the effectiveness of the feedback from the previous IAT in eliciting certain emotions. Feedback from the neutral IATs were not expected to be either strong positive or negative, feedback from prejudice-positive IATs were expected to elicit positive emotions such as pride, whereas feedback from the prejudice-negative IATs was expected to elicit guilt. Each IAT (including the actual task plus the feedback slides) lasted about 2 minutes, meaning that the entire IAT component of the experiment lasted less than 45 minutes. The entire experimental procedure including the questionnaire thus lasted about 60 minutes.

At the conclusion of the study, the researcher, in order to comply with ethical principles, thoroughly debriefed the participant to make sure that no harm was done. Participants were informed about the true nature of the study and were told that the prejudice feedback was pre-programmed and false. Additionally, it was explained to them that deception was necessary to ensure that the emotions they experienced were elicited naturally and on-line. This post-experimental interview assessed whether participants believed the prejudiced feedback. It, furthermore, contained a post-experimental emotion checklist where participants had to choose the one emotion they felt most strongly in connection with the feedback (see Appendix D).

RESULTS

Screening measures and participant selection

Descriptive statistics from the screening questionnaires (see Table 1) were used to derive cut-off scores that served as the inclusion criteria. Only participants who indicated availability for the follow-up study and who met the general demographic inclusion criteria (i.e. female, white, non-Jewish) were considered further. Out of the entire screening sample (N = 147) 59 individuals were selected based on their demographic details. Their responses on the screening questionnaires were analyzed and from the initial 59 individuals 19 participants were selected to participate in the laboratory experiment.

Table 1

Screening measures

	EMS Scores					Thermometer Rating			
	Black	Fat	Jewish	Gay	MCSDS	Black	Fat	Jewish	Gay
Selected participants n = 19									
Median	30.50	26.50	25.50	26.00	13.00	85.00	80.00	90.00	90.00
Mean	27.50	26.00	25.23	25.55	12.82	82.05	75.91	88.00	88.41
SD	8.46	6.05	9.31	7.94	4.26	13.86	18.17	7.58	12.85
Non-selected participants n =40									
Median	24.50	22.00	21.00	22.00	11.00	50.00	50.00	50.00	50.00
Mean	24.28	22.75	21.08	22.40	11.58	58.63	53.13	64.25	57.75
SD	6.59	6.57	6.65	8.13	3.40	15.97	19.04	18.80	19.93
Difference in means	3.23	3.25	4.15	3.15	1.24	23.42	22.78	23.75	30.66
Difference in medians	6.00	4.50	4.50	4.00	2.00	35.00	30.00	40.00	40.00

The rating thermometer was administered four times, once for each social prejudice category. Selected participants averaged higher across all four categories (M = 83.59, SD = 5.89) than those not included (M = 58.44, SD = 4.56).

The EMS scores were reviewed to identify participants with low prejudiced attitudes. Selected participants averaged higher across all four categories (M= 26.07, SD = 1.00) than those not included (M = 22.63, SD = 1.32).

Further analysis included the comparison of the results between participants selected for the actual experiment and those individuals not fulfilling the selection criteria. First the rating thermometer scores were considered where on average across all categories selected participants scored 25.15 degrees higher than those not selected. The difference in means and medians between the two groups for the total MCSDS score was 1.24 and 2.00 points, respectively. The average across all four categories for EMS was 3.44 points higher for selected participants than for those not selected.

Experimental data

A total of 147 students completed the online pre-screening survey. From this sample only 19 individuals met the inclusion criteria for the study. In 7 of the 19 cases the manipulated feedback for the critical IAT condition did not elicit feelings of guilt. Although the researcher stressed the reliability and accuracy of the IAT in determining implicit attitudes prior to commencing with the experiment the 7 participants did not believe feedback that stated that they are prejudiced towards black and homosexual people. Thorough questioning revealed that all of these individuals had cross-racial and/or homosexual affiliations such as having a black boyfriend/sister, being gay or having close gay friends. Thus, specific social circumstances affect the potential to experience guilt under the prejudice-paradigm.

Following the identification of these confounding social variables the data were analysed excluding the 7 identified cases. The sample size for the data analyses reported below was therefore n = 12.

The emotion checklist completed after feedback from each IAT required participants to rate the degree to which 8 different emotions were experienced.[2] Table 2 summarises the descriptive statistics for the final sample, and Figure 2 gives a graphic representation of the same data.

Considering the nature of each emotion contained in the emotion checklist the 7 emotions were collapsed into 3 self-conscious emotion categories: the first category consisted of guilt,

[2]This included the rating of a neutral feeling. The neutral rating was, however, excluded from analysis as the researchers found that rating neutrality seems unwarranted for as one cannot feel more or less neutral about something.

shame and embarrassment and was labelled as moral-negative emotions (Mor-Neg); the second category consisted of fear and anxiety and was labelled as basic- negative emotions (Bas-Neg); the third category consisted of pride and satisfaction and was labelled as moral-positive emotions (Mor-Pos). A single average score for each emotion category was generated for each participant yielding 3 scores per IAT.

Table 2

Descriptive Statistics for the final sample (n=12)

	Prejudice- positive condition		Prejudice- negative condition		Neutral condition	
	Weight IAT	Religion IAT	Race IAT	Sexuality IAT	Sport IAT	Facial hair IAT
Emotions						
Guilt	1.67	1.25	3.25	2.58	1.75	1.50
	(0.98)	(0.87)	(1.54)	(1.56)	(1.14)	(1.00)
Shame	1.67	1.08	2.75	2.33	1.50	1.50
	(1.23)	(0.29)	(1.42)	(1.07)	(0.67)	(0.90)
Embarrassment	1.58	1.08	3.17	2.50	1.50	1.58
	(0.67)	(0.29)	(1.19)	(1.31)	(0.90)	(1.00)
Fear	1.17	1.08	1.75	1.25	1.50	1.00
	(0.58)	(0.29)	(1.06)	(0.45)	(1.00)	(0.00)
Anxiety	1.33	1.33	2.08	1.83	2.42	1.33
	(0.49)	(0.65)	(1.24)	(0.83)	(1.51)	(0.89)
Pride	2.00	2.25	1.33	1.25	1.92	1.75
	(1.28)	(1.36)	(0.78)	(0.62)	(1.24)	(1.29)
Satisfaction	2.42	3.08	1.50	1.17	2.67	2.42
	(1.00)	(1.31)	(0.67)	(0.39)	(1.15)	(1.00)

Note. Means are presented with standard deviations in parentheses.

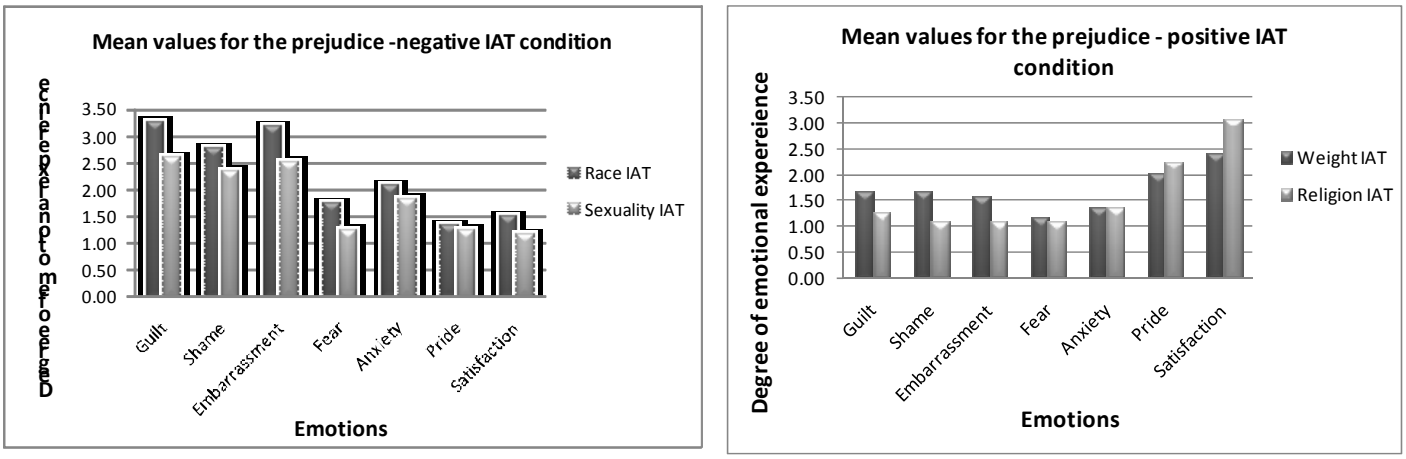


Figure 2. Bar chart of emotion means for prejudice-positive and prejudice-negative IAT conditions.

Two one-way repeated measures ANOVAs (see Table E2 & Table E5 in Appendix E) were used to analyse the emotion ratings for the prejudice-positive and the prejudice-negative IAT conditions. The cell mean plot (Figure 3) illustrates the pattern of the results.

A significant effect on emotion was found for the prejudice-positive IAT condition, $F(2, 36) = 21.1$, $p = 0.000001$, $\eta^2 = 0.64$. Post-hoc analysis of data from this IAT condition (see Table E3 in Appendix E) indicated that emotions in Mor-Pos category ($M = 2.44$, $SD = 1.02$) were elicited significantly more strongly than were emotions in the Mor-Neg ($M = 1.14$, $SD = 0.39$; $p = 0.00013$) and in the Bas-Neg ($M = 1.23$, $SD = 0.29$; $p = 0.00014$) categories. No statistically significant differences were detected between emotion elicitation in the Mor-Neg and Bas-Neg categories ($p = 0.94$). We can thus conclude that for the prejudice-positive conditions the manipulated feedback for the weight and religion IATs elicited pride and satisfaction most strongly, whereas the elicitation of guilt, shame, embarrassment, fear and anxiety were not significant.

A significant effect on emotion was found for the prejudice-negative IAT condition, $F(2, 22) = 9.8$, $p = 0.0009$, $\eta^2 = 0.47$. Post hoc analysis of data from this IAT condition (see Table E6 in Appendix E) indicated that emotions in the Mor-Neg ($M = 2.76$, $SD = 1.11$) category were elicited significantly more strongly than were emotions in the Mor-Pos ($M = 1.31$, $SD = 0.57$; $p = 0.0009$) and in the Bas-Neg ($M = 1.73$, $SD = 0.69$; $p = 0.02$) categories. No statistically significant differences were detected between emotion elicitation in the Bas-Neg and Mor-Pos categories ($p = 0.45$). We can thus conclude that for the prejudice-negative IAT conditions the

manipulated feedback for the race and sexuality IAT elicited guilt, shame and embarrassment most strongly, whereas the elicitation of pride, satisfaction, fear and anxiety was not significant.

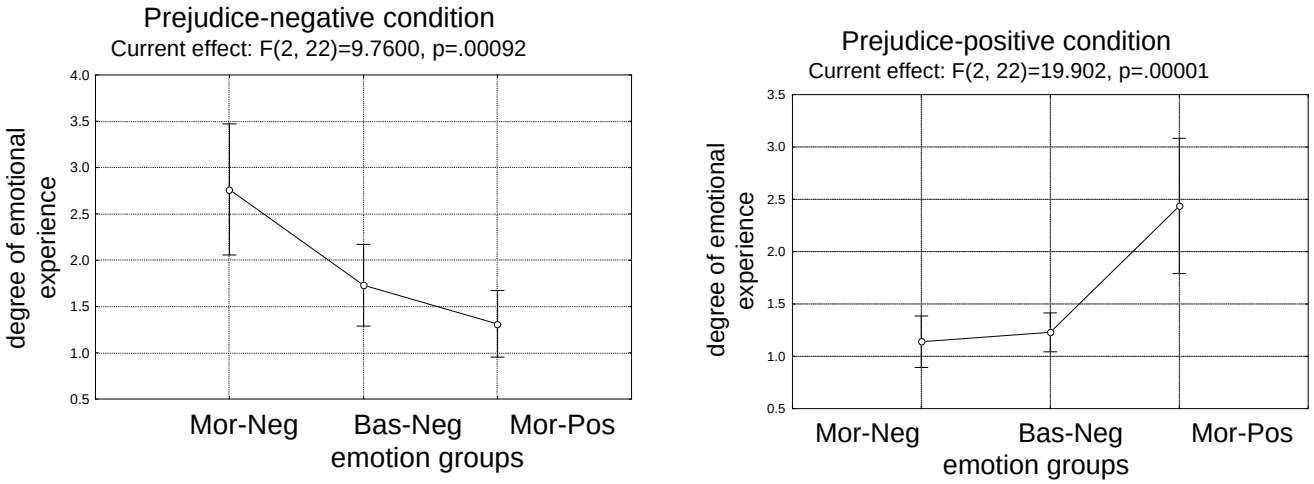


Figure 3. *Emotional experience during prejudice-positive & prejudice-negative conditions*

Comparison of means

A *t*-test for dependent samples was run on the Mor-Neg scores to determine whether the scores and therefore the degree of emotional experience is significantly different across the prejudice-positive and prejudice-negative IAT conditions (see Table E8 & Table E9 in Appendix E). This analysis aimed to support the findings above, and confirm that the experience of guilt, shame and embarrassment differed across the positive and negative feedback conditions. The analysis showed that the means of the Mor-Neg emotion group differed significantly across the two IAT conditions (M = 2.76, SD = 1.11 vs. M = 1.14, SD = 0.39), t = -4.78, p = 0.00058, r^2 = 0.68. These results confirm the effectiveness of the prejudice-negative IAT in exclusively eliciting the moral negative emotions reviewed.

A similar analysis was run to confirm that the Mor-Pos scores were statistically significantly different across the two IAT conditions. The *t*-test showed that the means of the Mor-Pos emotion group differed significantly across the two IAT conditions (M = 2.44, SD = 1.02 vs. M = 1.31, SD = 0.57), t = 4.03, p = 0.002, r^2 = 0.60. These results verify the effectiveness of the prejudice-positive IAT in exclusively eliciting the moral positive emotions reviewed.

Post-experiment emotions analysis

Upon completion of the IAT a post-experimental interview was conducted in order to identify the dominant emotion that was felt for the prejudice-positive and prejudice-negative IATs. For this purpose a further emotion checklist (see Appendix D) was completed. Here, rather than rating the intensity of 8 different emotions, participants were asked to choose the single emotion that they had most strongly experienced after each IAT. The initial analysis of this post-experiment emotion checklist lead to the exclusion of 7 biased cases, as previously explained. Further analysis of the post-hoc emotions revealed the following dominant emotions in the prejudice-negative IATs: for the race IAT 75% of participants experienced guilt (i.e. 9 out of 12; while 1 person felt shame, 1 felt neutral and another felt embarrassment). For the sexuality IAT 50% of the participants experienced guilt (i.e. 6 of 12; while 2 felt neutral, 2 felt embarrassment, 1 felt satisfied and 1 felt anger).

Post-experiment emotions analysis also revealed the following dominant emotions in the prejudice-positive IATs: for the weight IAT 67% of participants experienced satisfaction (i.e. 8 out of 12; while 4 felt neutral). For the religion IAT 75% of participants experienced satisfaction (i.e. 9 out of 12; while 3 felt neutral). No feelings of guilt, shame or embarrassment were elicited for the prejudice-positive IATs.

Investigation of potential relationships between emotion questionnaires and emotion elicitation

Correlations between the experimental questionnaire data and the collapsed emotion scores were generated to investigate whether relationships existed between participant's personal characteristics, as documented by the questionnaires, and the elicitation of certain emotions. For this purpose the following relationships were examined: the positive PANAS scores were correlated with emotion scores of the prejudice-positive condition; the negative PANAS scores were correlated with the emotion scores of the prejudice- negative condition; the AIM scores were correlated with the emotion scores of the prejudice- negative and prejudice-positive condition; the BIS scores were correlated with the emotion scores of the prejudice- negative and prejudice-positive condition; the BAS scores were correlated with the emotion scores of the prejudice- negative and prejudice-positive condition. For descriptive statistics of the questionnaire scores see Appendix E, Table E7.

Table 6

Correlation of emotion questionnaire scores & emotion elicitation during prejudice-positive condition

Correlations with prejudice- positive IAT condition			
	Mor-Neg	Bas-Neg	Mor-Pos
POS PANAS	-0.07	-0.18	-0.39
AIM	-0.15	-0.03	-0.33
BIS	0.33	0.33	0.73*
BAS Total	0.54	0.16	0.48
BAS Drive	0.43	0.06	0.12
BAS Fun	0.43	0.27	0.35
BAS Reward responsiveness	0.45	-0.01	0.73*

Note. * $p < 0.05$

Table 7

Correlation of emotion questionnaire scores & emotion elicitation during prejudice- negative condition

Correlations with prejudice- negative IAT condition			
	Mor-Neg	Bas-Neg	Mor-Pos
NEG PANAS	-0.48	-0.3	-0.17
AIM	-0.19	-0.16	-0.07
BIS	0.67*	0.43	-0.15
BAS	-0.13	-0.21	0.21
BAS Drive	-0.16	-0.01	-0.03
BAS Fun	-0.20	-0.19	0.31
BAS Reward responsiveness	0.09	-0.29	0.17

Note. * $p < 0.05$

The BDI-II scores were not considered in the analysis as they exclusively served as a control measure to account for depression as a confounding variable. Participants who according to the BDI-II scale would be categorised as depressed would be excluded from the data analysis. None of the 12 participants were rated as clinically depressed (M= 8.26, SD = 5.21) as all their scores remained below the score indicative of moderate to severe depression (20 and above).

Correlations between the PANAS scores and the emotion category scores (Mor-Neg, Bas-Neg & Mor -Pos) generated during either of the prejudice conditions were not significant (see Table 6 & Table 7). Similarly, no significant relationships were found between the AIM score and any of the emotion category scores (see Table 6 & Table 7).

A significant strong relationship was found between the BIS score and the Mor-Pos (r = 0.73) of the prejudice-positive condition as well as between the BIS score and the Mor- Neg (r = 0.67) emotion category of the prejudice-positive condition. Whereas the total BAS score yielded no significant relationships one of its sub-scores, *BAS Reward responsiveness,* revealed a strong positive relationship between the Mor- Pos (r = 0.73) emotion category of the prejudice-positive IAT condition.

DISCUSSION

The prejudice-paradigm proved effective in reliably eliciting guilt, shame, embarrassment and pride on-line by manipulating feedback on implicit prejudice measures. The current research also uncovered critical screening tools such as the consideration highly specific social and personality profiles that help to identify candidates that are particularly suitable for inclusion in an emotion elicitation study of this nature. Application of this direct elicitation paradigm will ultimately greatly strengthen and facilitate the investigation of moral emotions as the on-line elicitation of emotions will yield far more valid results than studies which use static images, or script driven imagery.

Many studies (Berthoz et al., 2006; Britton et al., 2006; Shin et al., 2000) have attempted to investigate moral emotions but all of these studies have been criticised in terms of the manner in which moral emotions were elicited and how that may have affected the research conditions and subsequent results. The challenge for research involving the elicitation of self-conscious emotions is to find a way to elicit the emotion on-line in an ecologically valid way within a laboratory setting. As a result, the current research devised an emotion-eliciting paradigm that was designed to elicit emotions in an on-line fashion by actively inducing certain target emotions in real time under very specific experimental conditions.

In the current emotion elicitation paradigm, manipulated IAT feedback aimed at eliciting moral negative and moral positive emotions. The design of the paradigm considered that emotions vary along a sociality dimension, where social dimensions such as guilt and pride are

evoked by self-reflection and self-evaluation (Eisenberg, 2000). Whereas guilt and shame are social negatively valenced emotions, pride falls within a social positive category. The key difference between the two categories of social emotions is, thus, whether self-evaluation follows a moral violation (negative valence) or socially valued outcome (positively valenced; Tangney, Stuewig & Mashek, 2006). Therefore the paradigm was constructed to create both situations of moral violation and moral conformity expected to yield negative and positive self-reflections, respectively. The analysis of the emotional experience as documented by the emotion checklist confirmed the predictions made by the researchers. The IATs were designed in such a way as to elicit moral negative emotions such as guilt during the prejudice-negative IATs and moral positive emotions such as pride during the prejudice-positive IATs. As intended, targeted emotions were elicited significantly more as compared to nontarget emotions during both IAT conditions, confirming the validity of the prejudice-paradigm. The type of emotion significantly elicited differed, depending on whether the prejudice-positive or prejudice-negative paradigm was administered. Moral negative emotions were elicited significantly more as compared to moral positive or basic negative emotions during the prejudice-negative IAT condition. Similarly, targeted moral-positive emotions were elicited significantly more as compared to moral negative and basic positive emotions during the prejudice-positive IAT condition. Finally, basic negative emotions were not significantly elicited during either of the IAT conditions confirming that the manipulation is targeting moral rather than basic emotional responses. These inferences were confirmed by the t-test analysis performed on the Mor-Neg and Mor-Pos emotion categories across both IAT conditions. Here it was shown that the Mor-Neg emotions were significantly elicited during the prejudice-negative IATs but not during the prejudice-positive IATs. Similarly, the Mor-Pos emotions were significantly elicited during the prejudice-positive IATs but not during the prejudice-negative IATs.

A great challenge in emotion elicitation research is the fact that certain moral emotions can co-occur in the same situation (Takahashi et al., 2004). For instance, in a moral transgression, a person could feel guilty about violating a social norm while feelings shameful about his/her own moral shortcoming (Eisenberg, 2000). A well known example thereof is the differentiation of shame and guilt as discussed by Eisenberg's paper (2000) on emotion regulation and moral development. In this study, we attempted to overcome this problem by treating emotions of the same nature, such as moral negative emotions, basic negative emotions

and moral positive emotions as related experiences. Although in the current research all emotions were measured separately for the analysis they were combined into three distinct emotional groups (Mor-Neg, Bas-Neg & Mor-Pos). The post-experimental checklist, however, aimed at ascertaining which specific emotion within the related emotion categories was elicited the most. Here, it was found that within the Mor-Neg emotion category guilt was elicited the most, whereas for the Mor-Pos emotion category satisfaction was elicited most frequently.

Further inferences can be made by correlating the experimental questionnaire scores and the emotion category scores. This analysis aimed at establishing a manner of identifying individuals, prior to the administration of the IAT, who will be most likely to experience a certain target emotion. The correlation between the BIS scores and the emotion categories revealed the most significant findings from the questionnaire data. Both Mor-Pos ($r = 0.73$) and Mor-Neg ($r = 0.67$) emotions were shown to have strong positive relationships with the BIS scores. This relationship suggests that a participant with a high BIS score is highly likely to experience pride and satisfaction during the prejudice-positive condition as well as experiencing guilt, shame and embarrassment during the prejudice-negative condition. Conversely, an individual scoring low on the BIS is less likely to experience pride and satisfaction during the prejudice-positive condition and also less likely to experience guilt, shame and embarrassment during the prejudice-negative condition. Whereas the total BAS score leads to no meaningful predictions the BAS Reward responsiveness score suggests a strong positive relationship between an individual's reward responsiveness and the Mor-Pos ($r = 0.73$) emotions on the prejudice-positive scale. Thus, a high BAS Reward responsiveness score indicates that a participant will experience pride and satisfaction on the prejudice-positive scale.

These findings are critical as the prejudice-positive IAT was designed to elicit pride and satisfaction, whereas the prejudice-negative IAT was designed to elicit guilt, shame and embarrassment. The results suggest that the BIS/BAS can identify individuals with suitable personal characteristics that increase the likelihood of specific target emotions being successfully elicited. The correlation analysis has shown that based on the BIS score the researcher can make a fairly strong prediction on whether a participant will experience the target emotions prior to the actual experiment. It should, therefore, be considered to include the BIS/BAS in the pre-screening selection process as the scale is indicative of a participant's suitability for the study.

The study also indicated that the participant inclusion criteria should be extended too. The current research stipulated that only white, female participants who are not Jewish, and who show low prejudiced attitudes and high social desirability should be admissible for the experimental follow-up study. The post-experimental interview has, however, clearly indicated that individuals involved in cross-racial or gay relationships should also be excluded, as these factors result in the participants not believing the manipulated feedback and hence not experiencing the target emotion. . It is therefore strongly suggested that the demographic information collected during the pre-screening should include indications of: sexual orientation; homosexual affiliations i.e. “Do you have gay friends?”; and cross-race contact i.e. Do you have black friends/boyfriend/family?”

The BDI-II should continue to be used but rather than forming part of the experimental questionnaires, it should be included in the pre-screening phase. This would allow exclusion of depressed individuals straight away. The pre-screening measures should thus ultimately aim at establishing an individual’s personality profile as the elicitation of emotions seems to depend on individual differences.

Overall the research revealed that the on-line elicitation of specific moral emotions using the prejudice-paradigm is effective but is highly dependent on participant selection. Manipulated IAT feedback proved effective as a manner of inducing feelings of guilt/shame/embarrassment as well as pride/satisfaction on-line given that the ‘right’ people were selected. Individuals are considered suitable if they meet the following profile: white; female; not Jewish; no black/homosexual affiliations; low prejudice; high social desirability; high BIS and high BAS Reward responsiveness score. For future application of the prejudice-paradigm it is imperative to follow the extended list of criteria to guarantee the effectiveness of the paradigm in eliciting specific target emotions on-line.

Limitations

One of the study’s limitations is the extensive and highly specific list of inclusion criteria. Due to this specificity of the participant profile it proved difficult to recruit suitable candidates as from a total of 147 pre-screened individuals only 19 participants were selected. As illustrated due to further social criteria not previously considered 7 further individuals were eliminated. Because the effectiveness of the paradigm heavily relies on the correct combination of personal

characteristics it is imperative to recruit individuals who fulfil all of the stipulated criteria. Therefore, while the success of the paradigm has been proven it remains a challenge to find suitable candidates who meet all of the identified criteria.

One-way RM ANOVAs were performed for the two critical IAT conditions (as presented in the results section above). The researchers considered running a MANOVA instead of the RM ANOVA as it does not assume compound symmetry. However, the MANOVA includes every possible pairwise comparison. For the purpose of the current research only certain specific comparisons were of interest, namely comparing the emotion elicitation scores across the two critical IAT conditions. This type of analysis could be accommodated by the 2 RM ANOVAs and the 2 t-tests. Thus, in order to focus the analysis on the planned comparisons of interest and under consideration of the inadequate sample size (n = 12) in relation to the experimental conditions (6 X 3) the RM ANOVA was chosen over the MANOVA.

Directions for future research

The common underlying problem within emotion elicitation research has been the lack of on-line elicitation of specific target emotions under laboratory conditions. The current research tested a prejudice-paradigm that aimed at actively eliciting emotions with the help of real-time moral digression rather than imagined moral violation. The development of such a novel and successful elicitation protocol required novel and innovative thinking that broke away from the usual repertoire of elicitation efforts seen in the literature. As the elicitation of moral emotions heavily relies on the sociality and morality of the situation (Eisenberg, 2000) a further challenge is to develop a protocol that is transferrable to a scanner environment and remains as reliable. As most social interactions cannot be replicated with a participant immobile and physically isolated in a scanner the current research has overcome this obstacle by developing a protocol that will remain as successful and reliable inside as outside of the scanner.

The current research served as a pilot study to test the effectiveness of the prejudice-paradigm in eliciting guilt on-line for a fMRI study aimed at investigating neural correlates of guilt elicitation. This brain imaging study requires the on-line and actual experience of guilt in a very specific moment i.e. when the participant is in the scanner. Considering the results above one may infer that the prejudice-paradigm proved successful in eliciting not only guilt but other

moral emotions such as pride and satisfaction in a controlled manner and can thus be adapted for use in the fMRI study.

REFERENCES

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt: implications for motivation and self-regulation in the context of prejudice. *Psychological Science, 18,* 524–530.

Barger, S. D. (2002). The Marlowe–Crowne affair: short forms, psychometric structure, and social desirability. *Journal of Personality Assessment, 79,* 286–305.

Beck, A. T., Steer, R. A., Brown, G. K. (1996). Manual for Beck Depression Inventory II (BDI-II). San Antonio, TX, Psychology Corporation.

Benedict, R. (1946). *The chrysanthemum and the sword*. Boston: Houghton Mifflin.

Berthoz, S., Gre`zes, J., Armony, J. L., Passingham, R. E., & Dolan R. J. (2006). Affective response to one's own moral violations. *NeuroImage, 31,* 945–950.

Britton, J. C., Phan, K. L., Taylor, S. F., Welsh, R. C., Berridge, K. C., & Liberzon, I. (2006). Neural correlates of social and behavioural emotions: An fMRI study. *NeuroImage, 31,* 397–409.

Cannon, W. B. (1927). The James-Lange theory of emotions: A critical examination and an alternate theory. *American Journal of Psychology, 39,* 106-124.

Carstensen, L. L., & Cone, J. D. (1983). Social desirability and the measurement of psychological well-being in elderly persons. *Journal of Gerontology, 38,* 713-745.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioural activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67,* 319-333.

Coan, J. A., Allen, J. J. B., & McKnight, P. E. (2006). A capability model of individual differences in frontal EEG asymmetry. *Biological Psychology, 72,* 198–207.

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43,* 245–265.

Crosby, E., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of black and white discrimination and prejudice: A literature review. *Psychological Bulletin, 87,* 546-563.

Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent psychopathology. *Journal of Consulting Psychology, 24,* 349-354.

Damasio, A. (1995). *Descartes' error: Emotion, reason, and the human brain.* New York: Avon.

DePaoli, L., & Sweeney, D. C. (2000). Further validation of the Positive and Negative Affect Schedule. *Journal of Social Behavior and Personality, 15,* 561-568.

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

Eisenberg, N. (2000). Emotion, regulation and moral development. *Annual Reviews Psychology, 51,* 665-697.

Ekman, P. (1999). Basic emotions. In T. Dalgleish & M. Power (Eds.). *Handbook of cognition and emotion* (pp. 45–60). Sussex, U. K.: Wiley.

Ferguson, T. J., Stegge, H., Damhuis, I. (1991). Children's understanding of guilt and shame. *Child Development, 62,* 827–839.

Ferguson, T. J., & Stegge, H. (1998). Measuring guilt in children: a rose by any other name still has thorns. In J. Bybee, (Ed.), *Guilt and children* (pp. 19–74). San Diego, CA: Academic.

Frijda, N. H. (1994). Universal antecedents exist, and are interesting. In P. Ekman & R. J. Davidson, (Eds.), *The nature of emotion: Fundamental questions* (pp. 155-162). New York: Oxford University Press.

Gaertner, S. L., & Dovidio, J. F (1986). The aversive form of racism. In X. E. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61-89). San Diego, CA: Academic Press.

Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice. *Journal of Experimental Social Psychology, 44*,164–172.

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

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

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3-25.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review, 108*, 814–834.

Hawkins, J. L. (1968). Associations between companionship, hostility, and marital satisfaction. *Journal of Marriage and the Family, 30*, 647-650.

Herek, G. M. (2000). Sexual prejudice and gender: Do heterosexuals' attitudes toward lesbians and gay men differ? *Journal of Social Issues, 56,* 251–266.

Johnson, S. L., Turner, R. J., & Iwata, N. (2003). BIS/BAS Levels and psychiatric disorder: An epidemiological study. *Journal of Psychopathology and Behavioral Assessment, 25,* 25-36.

Jorm, A. F., Christensen, H., Henderson, A. S., Jacomb, P. A., Korten, A. E., & Rodgers, B. (1999). Using the BIS/BAS scales to measure behavioural inhibition and behavioural activation: Factor structure, validity and norms in a large community sample. *Personality and Individual Differences, 26,* 49-58.

Keltner, D., & Buswell, B. N. (1996). Evidence for the distinctness of embarrassment, shame, and guilt: a study of recalled antecedents and facial expressions of emotion. *Cognition and Emotion, 10,* 155–71.

Kozma, A., & Stones, M. J. (1987). Social desirability in measures of subjective well-being: A systematic evaluation. *Journal of Gerontology,* 42, 56-59.

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., et al. (2000). Handedness and hemispheric dominance in healthy humans. *Brain, 123,* 2512 -2518.

Lange, C. (1885). One leuds beveegelser. In K. Dunlap (Ed.), *The emotions*. Baltimore: Williams & Wilkins.

Larsen, R. J. (1984). Theory and measurement of affect intensity as an individual difference characteristic. *Dissertation Abstracts International, 85,* 2297.

Larsen, R. J., & Diener, E. D. (1987). Affect intensity as an individual difference characteristic: A Review. *Journal of Research in Personality, 21,* 1-39.

Lewis, H. B. (1971). *Shame and guilt in neurosis*. New York: International Universities Press.

Lewis, M. (1992). *Shame: The exposed self*. New York: Free Press.

Lewis, M. (2000). Self-conscious emotions: embarrassment, pride, shame and guilt. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions*.(2nd ed., pp. 623-636). New York: Guilford.

Melvin, G. A., & Molloy, G. N. (2000). Some psychometric properties of the Positive and Negative Affect Schedule among Australian youth. *Psychological Reports, 86,* 1209-1212.

Millham, J. (1974). Two components of need for approval score and their relationships to cheating following success and failure. *Journal of Research in Psychology, 8,* 378-392.

Moilanen, D. L. (1993). Depressive information processing among non-clinical, non-referred, college students. *Journal of Counselling Psychology, 40,* 340–347.

Niedenthal, P. M., Tangney, J. P., & Gavanski, I. (1994). ‘‘If only I weren’t’’ versus ‘‘If only I hadn’t’’: distinguishing shame and guilt in counterfactual thinking. *Journal of Personality and Social Psychology, 67,* 584–595.

Nosek, B. A., Greenwald, A. G. & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin, 31*, 166 -180.

Panksepp, J. (1998). Affective neuroscience: The foundations of human and animal emotions. Oxford: Oxford University Press.

Panksepp, J. (2004). Affective consciousness: Core emotional feelings in animals and humans. *Consciousness and Cognition, 14,* 30–80.

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

Rusch, N., Holzer, A., Hermann, C., Schramm, E., Jacob, G. A., Bohus, M., et al. (2006). Self-stigma in women with borderline personality disorder and women with social phobia. *Journal of Nervous and Mental Disease, 194*, 766-773.

Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.) *Advances in experimental social psychology,* (pp. 49-79). New York: Academic Press.

Shields, S. A. (1991). Gender in the psychology of emotion: A selective research review. In K. T. Strongman (Ed.), *International review of studies on emotion* (pp. 227-245). New York: Wiley.

Shin, L. M., Dougherty, D. D., Orr, S. P., Pitman, R. K., Lasko, M., Macklin, M. L., et al. (2000). Activation of anterior paralimbic structures during guilt-related script-driven imagery. *Biological Psychiatry, 48,* 43-50.

Smith, R. H., Webster, J. M., Parrot, W. G., & Eyre, H. L. (2002). The role of public exposure in moral and nonmoral shame and guilt. *Journal of Personality and Social Psychology, 83,* 138–159.

StatSoft, Inc. (2007). STATISTICA version 8. Tulsa, OK.

Stones, M. J., & Kozma, A. (1994). The relationships of affect intensity to happiness. *Social Indicators Research, 31,* 159-173.

Takahashi, H., Yahata, N., Koeda, M., Matsuda,T., Asai, K., & Okubo, Y. (2004). Brain activation associated with evaluative processes of guilt and embarrassment: an fMRI study. *NeuroImage, 23,* 967– 974.

Tangney, J. P., Wagner, P., & Gramzow, R. (1992). Proneness to shame, proneness to guilt, and psychopathology. *Journal of Abnormal Psychology, 101,* 469–478.

Tangney, J. P., Wagner, P., Fletcher, C., & Gramzow, R. (1992). Shamed into anger? The relation of shame and guilt to anger and self-reported aggression. *Journal of Personality and Social Psychology, 62,* 669-675.

Tangney, J.P., Miller, R.S., Flicker, L., & Hill-Barlow, D. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology, 70,* 1256–1269.

Tangney, J. P., Niedenthal, P. M., Covert, M. V., & Hill-Barlow, D. (1998). Are shame and guilt related to distinct self-discrepancies? A test of Higgins 1987 hypotheses. *Journal of Personality and Social Psychology, 75,* 262-268.

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2006). Moral emotions and moral behaviour. *Annual Review of Psychology, 58,* 1–28.

Thompson, E. R. (2007). Positive and Negative Affect Schedule (PANAS): Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology, 38,* 227-242.

Tracy, J. L., & Robins, R. W. (2006). Appraisal antecedents of shame, guilt, and pride: support for a theoretical model. *Personality and Social Psychology Bulletin, 32,* 1339 -1351.

Watson, D., Clark, L. A., Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063-70.

Whisman, M. A., Perez, J. E., & Ramel, W. (2000). Factor structure of the Beck Depression Inventory second edition (BDI-II) in a student sample. *Journal of Clinical Psychology*, *56*, 545-551.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101–126.

APPENDIX A

IAT Description & Feedback Example

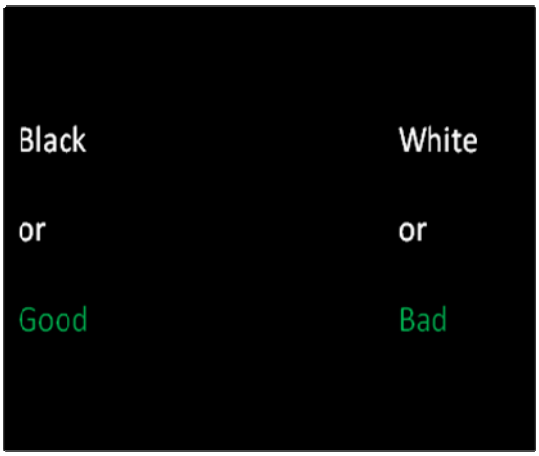
Participants were required to complete 6 IATs, each dealing with a different social prejudice category. In each IAT they were required to sort words or images into groups as quickly as they could while making as few mistakes as possible using the Q key (for categories on the left) and the P key (for categories on the right) on the keyboard. Participants were instructed that going too slow or making too many mistakes would result in an un-interpretable score, so they should try to work as quickly as possible while making as few mistakes as possible.

Words were either categorised as ‘good’ e.g. joy, love and pleasure, or as ‘bad’ e.g. agony, terrible, awful. Images were pictures of either black or white faces.

The process involved in the IAT is illustrated below using the *race* IAT as an example.

A screen would appear where categories were left and right aligned (see below).The ‘good’ and ‘bad’ words popped up on the screen one after the other and had to be assigned to the correct group using the keys. At the same time images of black or white faces would randomly appear on the screen which also needed to be assigned to the correct category.

The words and images were repeated twice during a sequence. Then the category alignments on the screen would change and a new sequence would begin where the same words and images had to be assigned to the new category configurations.



After the *race* IAT was completed participants were shown their results in the form of a brief feedback sentence.

Feedback

- Your Result: Your data suggest a strong automatic preference for white versus black people.

The brief feedback was followed by more elaborated feedback about their prejudice tendencies.

Elaborated Feedback

Generally, people with a strong automatic preference for white compared to black people:

- Are usually white themselves.
- Often find themselves in racially skewed working environments.
- Often react uncomfortably upon meeting new black people (e.g., by avoiding direct eye contact).
- Will avoid situations involving large crowds of black people.
- Will never openly disregard black people, but always prefer white company.
- Would never have a romantic relationship with a black person.

APPENDIX B

Emotion checklist

Please indicate the intensity with which you experienced each emotion below during the experiment? Please circle:

Neutral

Not at all Somewhat Very Much

1--------2--------3--------4---------5

Embarrassment

Not at all Somewhat Very Much

1--------2--------3--------4---------5

Pride

Not at all Somewhat Very Much

1--------2--------3--------4---------5

Guilt

Not at all Somewhat Very Much

1--------2--------3--------4---------5

Fear

Not at all Somewhat Very Much

1--------2--------3--------4---------5

Anxiety

Not at all Somewhat Very Much

1--------2--------3--------4---------5

Satisfaction

Not at all Somewhat Very Much

1--------2--------3--------4---------5

Shame

Not at all Somewhat Very Much

1--------2--------3--------4---------5

APPENDIX C

Cover Story & Consent Form

The literature shows that there is a great discrepancy between what people 'say' they feel (this is called 'explicit attitude') and what people 'really' feel (this is called 'implicit attitude') towards different social groups. We believe though that this implicit-explicit discrepancy has become smaller in recent years due to the educative value of the media and the different political times that children are born into. We are, thus, trying to prove that there is no longer a big difference between what people 'say' they feel and how they 'really' feel towards certain social groups.

I will briefly explain to you what you will be required to do during this study. As a first step you will be required to complete several questionnaires assessing some general aspects of your personality. The main body of the research study constitutes completing computer based tests, known as Implicit Association Tests (IATs). The IAT is an established test commonly used to measure implicit attitudes towards different social groups. (A large website is dedicated to it where people can complete a number of different versions of the IAT to assess their own implicit feelings towards certain social groups) Thus, with the IAT we will measure people's implicit attitudes (what they 'really' feel). Prior to starting with the study you will complete an example of an online IAT as practice.

You have been specifically selected for this study based on your explicit responses on the online survey you completed previously. These questionnaires indicated that you have positive attitudes towards most social groups. In this study we will now test your implicit (real) attitude towards these groups through the Implicit Association Test (IAT).

But don't worry, our previous tests have indicated that you are not a prejudiced person, so everything should be fine. Most people we've screened to participate so far did absolutely fine.

I will receive immediate feedback of your results while you are doing the tests.

UNIVERSITY OF CAPE TOWN
DEPARTMENT OF PSYCHOLOGY

Informed Consent to Participate in Research and Authorization for Collection, Use, and Disclosure Performance and of Other Personal Data

You are asked to participate in a research study conducted by Nadine Kilchenmann and Melike Nel (from the Psychology Department at the University of Cape Town). Your participation in this study is entirely voluntary. Please read the information below and ask questions about anything you do not understand, before deciding whether or not to participate.

1. Name of Participant

_____

2. Title of Research Study

Implicit and explicit attitudes toward different social groups among UCT students.

3. Contact Details

Main Investigator: Melike Nel

Department of Psychology
University of Cape Town
Office Number: 021 650 3415

4. Purpose of this research study

The literature shows that there is a great discrepancy between what people 'say' they feel (this is called 'explicit attitude') and what people 'really' feel (this is called 'implicit attitude') towards different social groups. We believe though that this implicit-explicit discrepancy has become smaller in recent years due to the educative value of the media and the different political times that children are born into.

We are, thus, trying to prove that there is no longer a big difference between what people 'say' they feel and how they 'really' feel towards certain social groups. And so far our data support this hypothesis. You have been specifically selected for this study based on your explicit responses on the online survey you completed previously. These questionnaires indicated that you have a positive attitude towards most social groups. In this study we will test your implicit (real) attitude towards these groups through the Implicit Association Test (IAT).

5. Procedure- what will be done if you take part in this research study

You will first be required to complete some questionnaires assessing some general aspects of your personality. The main body of the research study constitutes completing several computer-based tests, known as Implicit Association Tests (IATs). The IAT is an established test commonly used to measure implicit attitudes towards different social groups. You will have the opportunity to perform a practice version of this test to get used to it.

6. Length of study

The experiment consists of this one session only and should last about 1 hour. If at any time during the experimental sessions you find any of the procedures uncomfortable, you are free to discontinue your participation without penalty.

7. Potential risks and discomforts

During the study you should not experience any severe discomfort and no physiological/social/legal/financial risks are anticipated. You may experience slight fatigue during the experimental sessions. However, if you become tired you will be allowed to take breaks between blocks of tests whenever you want.

8. Credit for participation

By participating in this study you will receive 2 course credits for the Student Research Participation Programme (SRPP).

9. Can you withdraw from this research study?

You can choose whether or not to be in this study. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind or loss of benefits to which you are otherwise entitled. You may also refuse to answer any questions you do not want to answer. There is no penalty if you withdraw from the study and you will not lose any benefits to which you are otherwise entitled. If you withdraw information already collected may, however, still be used.

10. Confidentiality

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law. Confidentiality will be maintained in that you will be referred to as a number, not your name. The ‘key’ to match a number to a name will be kept separate from any obtained data. Furthermore, no personal details will be passed to anyone outside the research team.

11. Signatures

As a representative of this study, I have explained to the participant the purpose, the procedures and the possible risks of this research study.

_____ _____

Signature of Person Obtaining Consent & Authorization Date

You have been informed about this study's purpose, procedures and the possible risks of this research study. You have received a copy of this form. You have been given the opportunity to ask questions before you sign, and you have been told that you can ask other questions at any time.

You voluntarily agree to participate in this study. You hereby authorize the collection, use and sharing of your performance and other data. By signing this form, you are not waiving any of your legal rights.

_____ _____

Signature of Person Consenting & Authorizing Date

APPENDIX D

Post-experimental Interview & Post-experimental Emotion Checklist

Debrief/post-experimental interview

Participant:_____

Once all the IATs have been completed each participant will be thoroughly debriefed by the researcher where they will be informed about the true nature of the study and that the prejudice feedback was pre-programmed and false. It will also be explained why it was necessary to deceive them.

The debriefing session will take on the following form:

- How did you find the experiment:
  - were the instructions easy to understand
  - did feel like you could ask questions at any time during the experiment
  - did you find the experiment tiring
  - did you find the experiment too long

_____
_____
_____
_____

- Was any of the feedback you received to your responses different from what you would normally expect? If yes, for which IATs was this the case?

_____
_____
_____

- I saw the feedback you got for each IAT. How did the negative feedback make you feel? Were you surprised to be rated as prejudiced towards black and homosexual individuals?

_____
_____
_____

- Did you ever get the feeling that the experimenter might be interested in something else other than what was stated at the beginning of the study?

_____
_____
_____

- Well, actually we were not interested in implicit prejudiced attitudes. We manipulated the IAT feedback for the sexuality and race IAT. We pre-programmed it to indicate to you that you are prejudiced against these social groups no matter what your responses were. We did this because we wanted to test whether we can use this method to make people deliberately feel guilty. So these IAT gave no indication as to whether you are prejudice towards these people or not. So it does not represent your true prejudiced tendencies.

- Would you say that you felt guilty when you thought you were prejudiced against black and homosexual people?

_____
_____
_____

- Did you experience any other emotions e.g. anger or disbelief?

_____
_____
_____

- Did you feel that the feedback was genuine or did you suspect that it could be pre-programmed at any point? If so, how do think we could make seem more believable?

_____
_____
_____

- So we're sorry that we deceived you but if we would have told you that we are trying to make you feel guilty the experiment would not have worked. Do you have any questions?

_____
_____
_____

- It is crucial to our experiment that you please do not tell anyone about the true nature of the study.

Participant: _____

Emotion Checklist – Debriefing IAT

Please indicate the emotion you experienced the most after the *race* IAT, where the feedback indicated…

Neutral	**Embarrassment**
Pride	**Guilt**
Fear	**Anxiety**
Satisfaction	**Shame**

Please indicate the emotion you experienced the most after the *sexuality* IAT, where the feedback indicated…

Neutral	**Embarrassment**
Pride	**Guilt**
Fear	**Anxiety**
Satisfaction	**Shame**

Please indicate the emotion you experienced the most after the *weight* IAT, where the feedback indicated…

Neutral	**Embarrassment**
Pride	**Guilt**
Fear	**Anxiety**
Satisfaction	**Shame**

Please indicate the emotion you experienced the most after the *religion* IAT, where the feedback indicated…

Neutral	**Embarrassment**
Pride	**Guilt**
Fear	**Anxiety**
Satisfaction	**Shame**

APPENDIX E

Emotion Category Data Tables

Table E1

Descriptive Statistics of emotions under prejudice-positive condition

Descriptive Statistics			
	Mor- Neg	Bas- Neg	Mor-Pos
N	12	12	12
Mean	1.14	1.23	2.44
SD	0.39	0.29	1.02

Table E2

RM ANOVA summary table for emotions under prejudice-positive condition

Repeated Measures Analysis of Variance					
	SS	d.f.	MS	F	p
Intercept	138.19	1	138.19	156.80	0.000000
Error	15.86	18	0.88		
POS	15.33	2	7.67	21.1	0.000001
Error	13.10	36	0.36		

Table E3

Post-hoc summary table of emotions under prejudice-positive condition

Tukey HSD test Tests Error: Within MS = .36387, d.f. = 36.000			
POS	Mor- Neg	Bas- Neg	Mor -Pos
Mor - Neg		0.94	0.00013
Bas - Neg	0.94		0.00014
Mor - Pos	0.00013	0.00014	

Table E4

Descriptive Statistics of emotions under prejudice-negative condition

Descriptive Statistics			
	Mor- Neg	Bas- Neg	Mor -Pos
N	12	12	12
Mean	2.76	1.73	1.31
SD	1.11	0.69	0.57

Table E5

RM ANOVA summary table for emotions under prejudice-negative condition

Repeated Measures Analysis of Variance					
	SS	d.f.	MS	F	P
Intercept	134.82	1	134.82	201.68	0.00
Error	7.35	11	0.67		
NEG	13.40	2	6.70	9.76	0.00092
Error	15.11	22	0.69		

Table E6

Post-hoc summary table of emotions under prejudice-negative condition

Tukey HSD test Within MS = .68664, d.f. = 22.000			
NEG	Mor-Neg	Bas- Neg	Mor-Pos
Mor-Neg		0.02	0.00095
Bas-Neg	0.02		0.45
Mor-Pos	0.00095	0.45	

Pre-experimental Questionnaire Scores

Table E7

Descriptive Statistics of experimental questionnaire data

Descriptive Statistics			
	N	Mean	SD
NEG PANAS	12	23.53	40.82
POS PANAS	12	24.08	33.50
AIM	12	83.30	75.89
BIS	12	20.00	4.26
BAS Drive	12	11.42	1.98
BAS Fun	12	12.25	3.02
BAS Reward	12	16.83	2.04
BAS Total	12	40.50	5.76

Dependent t-test Analysis

Table E8

Comparing Mor-Neg emotion group across prejudice-positive and prejudice-negative condition

T-test for Dependent Samples (n=12)					
	Mean	SD	t	d.f.	p
Mor-Neg +	1.14	0.39			
Mor-Neg -	2.76	1.11	-4.78	11	0.00058

Table E9

Comparing Mor-Pos emotion group across prejudice-positive and prejudice-negative condition

T-test for Dependent Samples (n=12)					
	Mean	SD	t	d.f.	p
Soc-Pos +	2.44	1.02			
Soc-Pos -	1.31	0.57	4.03	11	0.002

AUTHOR NOTE

We thank Doug Mansfield, Department of Psychology, University of KwaZulu- Natal, for his critical assistance in the IAT programming.