NUI MAYNOOTH

Ollscoil na hÉireann Má Nuad

Developing the Function Acquisition Speed Test for Homonegativity

Aoife Cartwright, B.Sc. (Hons)

Thesis submitted in fulfilment of a Master of Science (M.Sc.) degree to the Department of Psychology, National University of Ireland, Maynooth.

October 2013

Supervised by Dr. Bryan Roche

Table of Contents

Declaration

I, the undersigned, hereby certify that this material, which I now submit in fulfilment of a M.Sc. degree, has not been previously submitted as an exercise for a degree at this or any other University, and is, unless otherwise stated, entirely my own work.

Signed: _____

Aoife Cartwright

Student ID: 67448716

Date: 30th October 2013

Acknowledgements

Firstly, I would like to sincerely thank my supervisor, Dr. Bryan Roche. Your constant advice, assistance, and endless encouragement has made this an incredibly fun learning experience. Thank you!

I would also like to thank everyone in the Psychology department at NUI, Maynooth, for always being there to help in any way they can.

I also have to extend my utmost gratitude to my wonderful family and friends. In particular, I would like to thank Claudia, Lisa, and Yasmin for their unending support.

Finally, I am especially grateful to my Mam and my partner Elliott – I literally couldn't have done this without you. Thank you for everything.

Abstract

Substantial research has been devoted in recent years to the development of a new and alternative test methodology to the standard explicit questionnaire (De Houwer, 2006). This current research was concerned with the laboratory development of one alternative methodology known as the Function Acquisition Speed Test (FAST; O'Reilly, Roche, Ruiz, Tyndall, & Gavin, 2012), into a functionally understood assessment of homonegativity, or anti-gay bias. In brief, the FAST was built around the idea that the stronger a pre-existing relation is between two classes of stimuli (e.g., Gay-Bad), the more resistance will be encountered when learning to form a new functional response class that is in opposition to it (e.g., Gay-Good).

This current research programme has focused primarily on establishing the optimal stimulus set for a Homonegativity FAST. Specifically, Experiments 1 and 2 utilised a known-groups approach to compare and contrast the relative effectiveness of semantically representative verbal and pictorial exemplars in a "single-phase" Gay-Good FAST. Experiments 3a and 3b investigated the suitability of more emotive pictorial stimuli high in affective valence in a "two phase" FAST, assessing the resistance to the formation of a Gay-Good class relative to neutral or contrasting functional stimulus classes.

Together, the findings from this research suggest that affective pictorial stimuli are the most suitable exemplars in a FAST for homonegative prejudice. Additionally, this research highlighted the enhanced stability of a "two-phase" FAST examining Gay-Good and Gay-Bad verbal histories simultaneously for providing a more accurate and conceptually valid assessment of homonegativity. Thus, over the course of this research, the FAST has been refined and developed into a functionally understood prototype for the assessment of homonegativity. More specifically, this research has identified an effective and representative

stimulus set, in addition to establishing a valid and stable two-phase FAST format capable of ascertaining an individual's homonegative verbal histories.

Chapter 1

Literature Review and General Introduction

1.1 Introduction

The current research is concerned with the development of a more reliable and valid alternative test format than the standard explicit questionnaire. It aims to add to the growing body of research devoted to the development of such measures and, more specifically, to present the Function Acquisition Speed Test (FAST) as a potentially reliable and valid psychometric tool. At present, there are two main schools of thought devoted to this particular research: Social Cognition, and Behaviour Analysis. As it is the mainstream approach, and by far the most well-known approach to developing such an alternative psychometric test format, the social-cognitive approach will first be reviewed, and a critical analysis of its theoretical foundations and current popular forms of measurement will be offered. This introduction will then turn its focus to the contrasting behaviour-analytic approach, and will present a review of the relevant literature in Behaviour Analysis devoted to developing an alternative test format to standard explicit measures. Finally, the recently developed behaviour-analytic measure known as the FAST will be presented, and the current research programme into the development of a FAST for measuring anti-gay prejudice will be introduced.

1.2 Explicit Measures

Traditionally, psychological researchers relied on self-report or explicit measures such as questionnaires, surveys, and interviews to assess or measure any given element of individual's personal history (e.g., De Houwer, 2003). That is, information regarding how an individual felt about or evaluated any given construct was ascertained by simply asking them directly via a survey or questionnaire. While this was the dominant method of data collection used by researchers to identify psychological information of interest for decades, there are a number of limitations inherent in the application of self-report measures to psychological

research (e.g., Greenwald & Banaji, 1995). The most significant of these limitations, social desirability bias and the 'introspection problem', will now be discussed.

1.2.1 Social Desirability Bias

One of the most prominent limitations of the use of standard self-report measures such as questionnaires in psychological research is social desirability bias (de Jong, Pieters, & Stremersch, 2012). Social desirability bias refers to the tendency of individuals to select as self-descriptive the response options for items that are considered more socially desirable than may necessarily be warranted by their corresponding traits or behaviours (Paunonen & Le Bel, 2012). In other words, social desirability bias refers to the tendency for people to respond on self-report measures in a way that is considered more socially favourable, or in a way that they believe may make a more favourable impression on the researcher (Paulhus, 1984; Paulhus & John, 1998). In particular, research indicates that when participants are asked questions that may be somewhat sensitive in nature, or which may refer to socially sensitive information (e.g., sexuality, race, gender, mental health, etc.), scores can be significantly distorted. For example, studies have found that, when asked directly, participants have intentionally distorted responses relating to mental health (e.g., Klassen, Hornstra, & Anderson, 1975), personality (see Furnham, 1986), sexuality and sexual practices (e.g., Fenton, Johnson, McManus, & Erens, 2001), and regarding socially-favourable activities such as voting (Holbrook & Krosnick, 2010).

1.2.2 The Introspection Problem

In addition to the issue of self-presentational distortions (i.e., social desirability), a limitation commonly encountered when utilising self-report explicit measures in psychological research is that individuals may not necessarily be consciously aware of the particular feelings or beliefs that are of interest to the researcher (Gawronski & De Houwer, in press). That is, while self-presentational concerns often distort information that

participants are *unwilling* to share, there are additional limitations to consider when using self-report measures relating to information participants may be *unable* to share. For instance, there are psychological attributes that may be of interest to the researcher, but that remain introspectively inaccessible or outside of the individual's conscious awareness. Simply put, participants may not be strongly aware that they hold a particular attitude and, thus, may not be able to reflect on and express it accurately on a self-report questionnaire (Dambrun and Guimond, 2004). It has long been argued that our ability to accurately introspectively access the psychological processes that guide our own behaviour is quite limited (e.g., Nisbett & Wilson, 1977; Egloff, Wilhelm, Neubauer, Mauss &, Gross, 2002), an issue that has been found to be particularly pronounced in socially-sensitive social or clinical psychological research. For instance, it has been found that, when asked directly, participants may be unable to accurately recount their privately held beliefs regarding socially sensitive attitudes toward race, sexuality or mental health (e.g., Banaji, 2001; Jost, Pelham, & Carvallo, 2002; Rudman, Greenwald, Mellott, & Schwartz, 1999; Spalding & Hardin, 1999; Wilson, 2002). Explicit self-report measures such as questionnaires, therefore, may be inherently unsuitable for this type of information (e.g., Cullen & Barnes-Holmes, 2008; Greenwald & Banaji, 1995).

1.3 Implicit Measures

Considering the above findings, it would appear that traditional explicit measures are neither an appropriate nor effective means of data collection in psychological research concerned with sensitive information. Consequently, there has been significant research devoted to the development of a more reliable and valid test format than the standard explicit measures, and into identifying methods of data collection that elicit more accurate information about privately held beliefs or behavioural intentions (De Houwer, 2003). While considerable research has been devoted to accurate means of data collection since the

establishment of psychology as a distinct science, particularly significant advancements have been made in this regard in recent decades, with the development of a number of alternative test formats that can be collectively termed *implicit measures*. Implicit (or indirect) measures are a group of psychometric-style tests that were developed in an attempt to address some of the limitations of explicit measures discussed above. While there are considerable methodological variations between different implicit measures, what distinguishes them as a group from the traditional self-report forms of assessment is that they each infer mental contents from participants' performance on some experimental paradigm rather than a self-reports (Gawronski, 2009).

This idea that psychological information of interest could potentially be accessed via performance on some task may date back to certain influential test formats developed in the 19th and 20th centuries. One of the most significant influences on the recently developed implicit measures arose from research conducted by Donders in the mid 19th Century, which indicated that the time required to perform some mental computation may reveal something significant about the processes of the mind (Lane, Banaji, Nosek & Greenwald, 2007). Donders reported that when mental tasks or computations were easier or more simplistic, the response times were lower with fewer errors on behalf of the participant, relative to more difficult or complex mental tasks (Donders, 1868, 1969). This led to the widely influential notion that an individual's mental states could potentially be inferred on the basis of their performance on some experimental task, as opposed to more fallible self-report measures.

One of the first test formats to incorporate this finding into a psychometric methodology was the Thematic Apperception Test (TAT; Morgan & Murray, 1935), a popular psychoanalytic 'projective' measure that infers information regarding an individual's various psychological attributes based on their responses to ambiguous images. In the TAT, an individual's patterns of thought, attitudes, observational capacity and emotional responses

regarding themselves, others and their environment are ascertained from their responses to a set of cards that portray human figures in a variety of settings and situations. Participants are required to tell the researcher a story about each of the test cards in terms of the event shown in the picture; what has led up to it; what the characters in the picture are feeling and thinking; and the outcome of the event. The information received from the participant is then used to infer information regarding some psychological attribute of interest.

Another important precursor to current implicit measures was the neuropsychological measure known as the Stroop Task (Stroop, 1935). The Stroop Task assesses attentional biases in humans through examining reaction time differences between 'consistent' and 'inconsistent' tasks. In the original Stroop Task, words depicting the names of colours are presented to participants who are asked to name aloud the colour in which the word is set while ignoring the actual typed word. For example, if the word "blue" was presented to participants in the colour yellow, the correct response would be to say "yellow" rather than "blue". What is referred to as the Stroop Effect is the finding that reaction times are consistently faster for the consistent (when the colour of the word matches the typed word) versus inconsistent (when the colour of the word does not match the typed word) tasks (Stroop, 1935). The original Stroop has been modified several times and has led to the development of a number of popular neuropsychological and clinical measures, such as the Emotional Stroop, a task that compares the response times of participants in naming colours of words with high negative emotional valence versus words with neutral emotional valence (see Williams, Mathews, & MacLeod, 1996, for a review).

The above test formats laid the groundwork for creating a test format for use in psychological research that could potentially overcome the previously discussed limitations of explicit measures. That is, they each supported the idea that information regarding an individual's inner state could potentially be inferred from their performance on an

experimental task or paradigm (thus overcoming the problem of introspection) wherein the nature of the task is not overtly discernible to the participant (thus overcoming the problem of social desirability distortions). In this way, these measures have been particularly influential for more recent research conducted within the respective fields of Social Cognition and Behaviour Analysis, from within which several new test formats have emerged in recent years.

1.4 Social Cognition and Implicit Test Development

The social-cognitive approach to implicit testing expanded upon the idea that an individual's inner world could be assessed and measured via their performance on an experimental paradigm and extended it to what is known in Social Cognition as the 'attitude' concept. More specifically, implicit measures that have emerged from within the field of Social Cognition were developed as a way to improve the psychometric tests available to psychological researchers by incorporating findings from research investigating what are known as 'implicit attitudes' into the development of a new and indirect psychometric test format (e.g., De Houwer, 2003; Gawronski, 2009).

In Cognitive Psychology, and more specifically in the field of Social Cognition, one of the most dominating topics of the last century has undoubtedly been the attitude concept. Declared in 1935 by Allport to be "probably the most distinctive and indispensable concept" (p. 198) in Social Psychology, the concept of attitudes has captivated researchers in the field of Social Cognition for decades, and is central to the social-cognitive approach to Social Psychology (Collier, Minton, & Reynolds, 1991). Thus, research examining how exactly attitudes should be measured has dominated the literature in recent years, with hundreds of studies annually devoted to effective and accurate attitude measurement. A review of the relevant literature of the social-cognitive approach to attitudes will be presented below, along

with a review of some of the most popular methods used in Social Cognition to measure socially relevant attitudes.

1.4.1 The Attitude Concept

Hypothetical constructs, mentalistic concepts, and conceptual analogies are frequently employed in Cognitive Psychology and the sub-discipline of Social Cognition to conceptualise how information is processed, analysed, and stored in the brain (e.g., De Houwer, 2011; Kruglanski & Stroebe, 2012). The notion of an 'attitude' is one such example of a hypothetical construct that is used to conceptualise a proposed actual existing cognitive or mental construct. While there is no universally agreed-upon definition for what exactly is represented by an attitude, generally in the field of Social Cognition, an attitude is proposed to represent an evaluative *cognitive association* (e.g., between a concept and an attribute) that has been established in experiential history of an individual towards any given element of their environment, which will predispose a favourable or unfavourable behavioural response toward that element (see Olsen & Zanna, 1993, for a review). Thus, the term 'attitude' is an umbrella term of sorts conceptualising all evaluations held by individuals, including self-esteem, stereotypes, biases, general negative or positive evaluations, etc., and is one of the most widely accepted concepts used in Social Cognition to account for what is known is the field as the 'cognitive mediation' of social behaviour (Greenwald & Banaji, 1995).

1.4.2 Implicit Attitudes

A number of prominent social-cognitive researchers have additionally proposed that there are different, quite distinct types of attitudes, which are used to conceptualise the difference between evaluations made by individuals that are conscious and direct in nature versus those that are indirect or unconscious in nature (see Greenwald & Banaji, 1995). This distinction has had particular influence over how attitudes are examined in Social Psychology, and has directly influenced modern-day attitude measurement. Traditional

means of examining an individual's attitude in social psychological research utilised self-report measures such as questionnaires, surveys, and interviews (e.g., De Houwer, 2003; Gawronski, 2009). However, as discussed above, while the use of explicit measures might be useful when the research is interested in examining attitudes toward certain types of information, they have proven problematic when investigating socially sensitive information such as individual attitudes toward race, gender, sexuality, etc. (e.g., Cullen & Barnes-Holmes, 2008).

A number of explanations have been offered from within the paradigm of Social Cognition as to why this is the case. Greenwald and Banaji (1995), for example, suggested that traditional explicit psychometric measures were unsuited to the measurement and assessment of socially sensitive attitudes for two reasons. Firstly, the authors suggest that explicit measures are not psychometrically suitable for assessing these types of attitudes due to the previously discussed issue of social desirability biases (i.e., concealing your 'real' attitudes from researchers because it may seem socially undesirable), which they argue is compounded by the low procedural implicitness of direct measures (i.e., the level to which the nature of the test is discernible to the participant). Secondly, Greenwald and Banaji (1995) suggested these types of measures are not suited to the assessment of this kind of information as they propose that social behaviour is mediated more by implicit attitudes that may potentially be unaware to the individual and, therefore, cannot be adequately or accurately assessed by explicit measures. By nature, according to the authors, social cognitions and behaviours are implicit, unconscious and indirect, and are thusly inaccessible to traditional explicit measures.

Expanding on the idea that Social Cognition is mediated more by implicit rather than explicit attitudes, Greenwald & Banaji (1995), proposed a new theory of social behaviour they refer to as 'Implicit Social Cognition'. Implicit Social Cognition is described as the

indirect, unconscious, or implicit mode of operation for attitudes and stereotypes, and is centred around the concept of implicit attitudes, defined as "the introspectively unidentified (or inaccurately identified) trace of past experience that mediates [favourable or unfavourable feeling, thought, or action toward social objects]" (Greenwald & Banaji, 1995, p. 8). In other words, it refers to how implicit cognitions are involved with, and how they interfere with, deliberate judgements. In other words, while attitudes are generally assumed to reflect cognitive associations between a concept and an attribute, implicit attitudes are proposed to reflect the associations that are outside of direct conscious awareness. In light of this theory, Greenwald & Banaji highlighted the need for a reliable laboratory measure that enables the efficient assessment of individual differences in implicit social cognition, or implicit attitudes.

1.5 Implicit Attitude Measurement

A number of such measures have emerged from within the field of Social Cognition that claim to assess implicit cognitive associations between concepts and thus the underlying implicit attitudes. The most popular of these 'associative' implicit tests are the Implicit Association Test (Greenwald, McGhee & Schwartz, 1998), the Evaluative Priming Task (Fazio, Jackson, Dunton & Williams, 1995), the Semantic Priming Task (Wittenbrink, Judd & Park, 1997), the Go/No-Go Association Task (Nosek & Banaji, 2001) and the Extrinsic Affective Simon Task (De Houwer, 2003). These tests, though they differ methodologically, have in common certain methodological features, in addition to similar theoretical foundations.

Essentially, what is common across these measures is that they each rely on the aforementioned assumption that social attitudes are somewhat associative in nature (i.e., that implicit attitudes reflect unconscious cognitive associations between different mental representations; Hughes, Barnes-Holmes & De Houwer, 2011). Hughes and colleagues posit

that throughout the Implicit Social Cognition literature, the dominant 'associative' assumption underlying theoretical accounts of implicit attitudes has largely influenced each of the major techniques designed to measure them. That is, each of the main social-cognitive implicit measures was developed based on the associative idea that social attitudes could be inferred from a methodology that accesses pre-existing cognitive associations. This associative assumption underlying these methods is the predominant way in which the tests listed above have been understood and their effects interpreted (De Houwer, 2011; Hughes et al., 2011). De Houwer and De Bruycker (2007), for example, reiterate the common idea in social-cognitive literature that attitudes are assumed to represent stable, trait-like cognitive associations (e.g., a Black person–bad association) between 'positive' and 'negative' and the concept of interest. Thus, the social-cognitive measures share a common methodological foundation in that each aims to assess whether an individual feels favourably or unfavourably toward something (e.g., a particular social group) by measuring the 'associative strength' between the category of interest and a positive or negative attribute (e.g., Hofmann, Gawronski, Gschwender, Le & Schmitt, 2005).

1.5.1 The Implicit Association Test

Of these associative measures, the most popular measure by far has been the Implicit Association Test (IAT; Greenwald et al., 1998). Building upon their theory of how implicit cognitive associations established in the individual's history mediate social feelings, thoughts, and actions, Greenwald and colleagues proposed a test that claims to measure the strength of socially significant cognitive associations or associative structures. As mentioned above, these associative structures are believed to underlie implicit attitudes, which in turn mediate behaviours and judgements, and thus the relative strength of the associations should theoretically indicate the strength of the implicit attitude. In other words, if there is a strong association between an image of an insect and the word 'bad' as opposed to the word 'good',

and a weak association between an image of a flower and the word ‘bad’ as opposed to the word ‘good’, then the IAT is proposed to have revealed the pre-existing association that would underlie the individual’s implicit attitude towards insects and flowers (Greenwald et al., 1998).

More specifically, the IAT is a computer-based task that consists of a series of phases, or blocks, containing a fixed number of trials, within which different stimuli and attribute terms are presented on-screen (e.g., Greenwald et al., 1998; Greenwald, Nosek, & Banaji, 2003). Typical IATs begin by presenting participants with the target categories followed by the attribute concept. The first phase of the IAT usually involves training participants to press a response key located to the left of the computer keyboard for one target concept (e.g., insects) and to press a response key located to the right of the computer keyboard for the other concepts (e.g., flowers). The second phase of the IAT employs the same structure, however, participants are then trained to press left and right response keys for the two target attribute categories (e.g., pleasant and unpleasant). The third and fourth phases, presented in a random order, then require participants to respond to both a target stimulus and an attribute stimulus concurrently by pressing a left key press for one combination (e.g., flowers + good) and a right key press for the other combination (e.g., insects + bad). The stimuli and attributes are presented within one block in a manner that is proposed to be consistent with a pre-existing implicit association (e.g., flowers-good and insects-bad) and within another in a manner proposed to be inconsistent (e.g., flowers-bad and insects-good).

The strength of these associations is then measured in terms of response latencies for each computer-administered categorisation task, such that if a person finds one more difficult (evidenced by longer response latencies) and finds another comparatively easy (evidenced by shorter response latencies), it is inferred that there is a stronger association between the paired stimuli presented in the easy task (Lane et al., 2007). The relative strength of

association is then proposed to reflect relative implicit preference for the pairing in the easier consistent task (e.g., flower-good/insect-bad) over the pairing in the more difficult inconsistent task (e.g., flower-bad/insect-good; Lane et al., 2007), with results suggesting that responses are faster when associated categories are assigned to the same response key on a response recording device than when they are assigned to different keys. Thus, when attitudes are conceptualised as cognitive associations between the target concept and the concepts 'positive' or 'negative', the IAT (and similar measures) can be utilised as a means to assess socially relevant implicit attitudes (De Houwer & De Bruycker, 2007).

Greenwald and colleagues suggest that, in this way, the IAT may resist masking by self-preservation strategies, and may preclude the concealment of attitudes that an individual may either intentionally wish to conceal, or may not necessarily know they hold (as they are largely unconscious). The IAT is, therefore, one of the primary ways in which social-cognitive researchers suggest circumventing the problems inherent in self-report measures. Since its publication in 1998, the IAT has received significant research attention, with over 200 papers reporting the use of the method. Given that it is believed to be superior to many explicit attitude tests in terms of being capable of overcoming experimental demand characteristics and precluding faking responses (see Banse, Seise, & Zerbes, 2001; Kim, 2003), the IAT has become a widely used measure not only in social and clinical research, but also in a variety of other contexts. For example, it has been used to measure implicit attitudes in socially sensitive domains such as race (Baron & Banaji, 2006; Greenwald et al., 1998; Greenwald, Oakes, & Hoffman, 2003; Nosek, 2005) and gender (Aidman & Carroll, 2003; Greenwald & Farnham, 2000; Nosek, 2005; van Well, Kolk, & Oei, 2007), as well as in clinical (Egloff & Schmukle, 2002), developmental (Baron & Banaji, 2006), and health settings (Czopp & Monteith, 2003). Its use has also been extended to fields as diverse as marketing (Maison, Greenwald, & Bruin, 2001) and legal scholarship (Kang & Banaji, 2006).

1.6 Conceptual and Methodological Issues with the Implicit Association Test

While the IAT is well established as an approach to attitude measurement, numerous concerns have been raised regarding its validity as a psychometric measure. A comprehensive review of these concerns would be beyond the scope and practical constraints of this thesis, and indeed it should be noted that the aim of the current research is not to evaluate social-cognitive measures such as the IAT. However, some of the important methodological and conceptual critiques of the IAT were borne in mind when developing the measure known as the Function Acquisition Speed Test, which is the focus of the current thesis. Therefore, a brief summary of some of the main unresolved issues in the IAT relevant to the current research of these will be briefly outlined below.

One of the most significant criticisms of the IAT is that research into the core processes underlying the IAT is lacking despite widespread adoption of the IAT in a range of contexts (Fazio & Olsen, 2003). While the IAT is generally assumed to assess an individual's implicit attitudes via association strengths, there remains considerable confusion regarding what precisely is being measured, and what exactly accounts for the IAT effect (e.g., Rothermund & Wentura, 2004). More specifically, there is considerable conceptual confusion surrounding how the task structure in the IAT actually determines implicit association strength, with particular confusion surrounding the poorly defined mentalistic association concept itself (Hughes et al., 2011). For instance, the empirical support for Greenwald and Banaji's (1995) theory of 'associative' implicit attitudes has come primarily from the IAT effect, which itself relies on the associative assumption within the implicit cognition theory for explanation (Greenwald et al., 1998). Indeed, a review by Fiedler, Messner and Bluemke (2006) reported the clear lack of any unanimously agreed upon testable model that can account for the IAT effect. Furthermore, it has been suggested that the explanations that have been provided for the test's effect by IAT researchers may actually

rest on inherently un-testable assumptions (e.g., Roche, O'Reilly, Gavin, Ruiz, & Aranciba, 2012).

It has also been suggested that the lack of any unified model explaining the IAT effect is further compounded by the fact that scores on current forms of the IAT are confounded by numerous influences. That is, research has revealed that the IAT effect may be modulated by extraneous variables unrelated to the underlying implicit attitude. The stimuli used to represent the target categories, for instance, have been found to exert dramatic influence on both the magnitude and the direction of the IAT effect (Bluemke & Friese, 2006; Govan & Williams, 2004; Rothermund, Wentura & De Houwer, 2005; Steffens & Plewe, 2001). Additionally, a range of interpersonal variables have also been found to contribute to the IAT effect, such as cognitive capacity (McFarland & Crouch, 2002; Mierke & Klauer, 2003), previous experience with similar implicit measures (Fiedler & Bluemke, 2005) or a fear of appearing prejudiced (Frantz, Cuddy, Burnett, Ray & Hart, 2004; Vanman, Saltz, Nathan & Warren, 2004).

Additionally, several researchers have also criticised the core assumption that the IAT effect is representative of implicit attitudes, given the fact that the IAT can only provide an index of relative rather than absolute association strength and, as such, cannot be used to measure the valence of individual concepts (e.g., Blanton, Jaccard, Gonzales, & Christie, 2006; De Houwer, 2002). For example, an individual who is assumed due to longer response latencies on a Black-good/White-bad association task compared to a Black-bad/White-good association task to be racially biased toward Black people, could merely be positive towards Black people, but even more positive toward White. Thus, rather than providing a score that represents an individual's bias, it may only represent a preference for one stimulus category over another, which may, in this way, be somewhat arbitrary.

Additionally, several researchers have raised concerns regarding the IAT's poorly understood scoring practices (see Arkes and Tetlock, 2004; Blanton and Jaccard, 2006, 2008; Tetlock and Mitchell, 2009). Firstly, considerable concerns have been raised generally regarding the wide use of reaction times as an assessment criterion within implicit measures. Within the IAT, the response times are typically measured from the point of stimulus presentation to the end of the trial (i.e., where a correct response is pressed), for both the correct response and the incorrect response. However, corrective feedback (i.e., wrong) is only presented after an incorrect response, after which participants are required to respond correctly. Therefore, the response times for incorrect trials are considerably larger than for correct trials, given that they contain the time taken for error feedback to be presented and the alternative correct response to be made. This is intentional on the part of the developers of the IAT, and is designed to negate the need to impose an arbitrary time penalty added to recorded response times, a practice formerly used by IAT researchers to artificially increase reaction times for trials on which an error was made (Gavin, Roche, Ruiz, Hogan, & O'Reilly, 2012). This was done in order to "fairly" reflect processing time that should be longer when errors are made compared to when they are not. The idea is that even when processing time are short during an incorrect response, this is due to lack of 'correct' processing, and had this occurred it would have taken longer for the correct response to be produced. Put simply, a review of the literature on the IAT scoring practices indicates that the researcher simply estimates at how long it would have taken to produce that correct response, deletes the actual reaction time from their data sheet and replaces it with an artificial "correct" one. The artificial reaction time inserted for trials on which an error was observed is usually the average response time for trials as a whole plus a constant of around several hundred milliseconds. However, it is more common now to use the error feedback procedure in order to build the time penalty in to recorded reaction times, thus obscuring it from view

and burying it in an algorithmic computation (see Greenwald, Banajii & Nosek, 2003). The IAT relies on a method in which the raw IAT scores (i.e., participant error rates and reaction times) are transformed using what is known as the *D* algorithm (see Glashouwer, Smulders, de Jong, Roefs, & Wiers, 2013 for a recent review) to create a hybrid IAT score. Behaviour analysts have argued that the use of feedback following incorrect responses only, and the requirement to effectively make an observation response only following errors, creates an imbalance in terms of response times, whereby an increased error rate results in considerably larger (partially artificial) response latencies, which in turn produces an exaggerated IAT effect (i.e., longer response latencies for inconsistent rather than consistent trials in line with experimental hypotheses; Gavin et al., 2012). However, some researchers have objected to this statistical technique, arguing that does not accurately reflect either response rates or response time (see Gavin et al., 2008, for a full review). In contrast, these researchers have suggested that response accuracy or indeed fluency (combined accuracy and speed) rather than speed alone might be superior as an assessment of task difficulty within implicit measures (Gavin et al., 2012; O'Reilly, Roche, Ruiz, Tyndall, & Gavin, 2012).

The use of normed scores within the IAT is an additional point of contention in the literature. The IAT typically standardizes raw scores to give a normed value, which is perceived to be representative of the strength of association (and thereby of attitude strength) between the underlying construct and the attribute (i.e., 'strong', 'moderate' or 'weak'; Blanton & Jaccard, 2006). That is, the responses of a single participant across multiple trials are normed to create a standardised IAT score, which is then assumed to reflect a level of bias. However, as argued by Blanton and Jaccard (2006), this normed score is in fact no more representative of underlying prejudice or preference than raw IAT data in the form of response latencies. This is largely due to the fact that no empirical studies have, as of yet, quantified IAT scores in terms of their absolute strength (i.e., to quantify and standardise

scores against associations of known strength) and consequently the used of normed scores may be considered somewhat arbitrary.

Lastly, criticisms have also been raised regarding the procedural implicitness of the IAT (i.e., the extent to which the aim of the test is discernible to the participant), which has been suggested as one of the primary ways the IAT circumvents the issues in explicit measures (Greenwald et al., 1998). Some researchers have proposed that the IAT effect is heavily susceptible to faking, that is, it may be possible, even with the indirect format outlined above, to fake results on the IAT (e.g., Cvencek, Greenwald, Brown, Gray, & Snowden, 2010; Steffens, 2004). This effect is found to be particularly pronounced when participants have had previous experience with the IAT procedure (Fiedler & Bluemke, 2005). Further to this, it has been suggested that the IAT's procedural implicitness is somewhat compromised by the fact that the association under analysis (e.g., Black-Bad) is easily discernible to the participant, even though the specific purpose of the test is not overtly obvious (Roche et al., 2012).

1.7 Behaviour Analysis and Implicit Test Development

The second school of thought dedicated to the development of a more reliable and valid test format to the standard explicit questionnaire is Behaviour Analysis. A number of tests have been developed from within this field in recent decades, which may provide a valid alternative to the mainstream Social Cognition approach (Hughes et al., 2011). Incidentally, based on the critiques outlined above, a more functional approach has in fact been called for in recent years by prominent social-cognitive researchers, as a means to understand the social-cognitive implicit measures in functionally understood terms (e.g., De Houwer, 2011). Before outlining the modern behaviour-analytic approach to implicit testing, however, it is firstly necessary to introduce the paradigm from within which these tests have been

developed. Consequently, the behaviour-analytic concept of derived relational responding will be briefly reviewed.

1.7.1 **Derived Relational Responding**

The functional analysis of verbal behaviour is one of the defining features of a behaviour-analytic approach to human language and cognition (Leigland, 1997). Therefore, one issue that is particularly pertinent in modern Behaviour Analysis is the concept of 'Stimulus Equivalence' (Hayes & Wilson, 1993; Sidman, 1971, 1994). Stimulus Equivalence refers to the ability to relate certain stimuli in the absence of explicit reinforcement of that relation, such that knowing that if A is the same as B, and B is the same as C, then A and C must be the same, or related. That is, within an experimental paradigm whereby choosing A when presented with B was reinforced, and choosing C in the presence of B was reinforced, A and C would each acquire discriminative functions (i.e., it would increase the likelihood) for choosing B, but, in addition, A and C would each acquire discriminative functions for choosing the other, without the two being in any directly trained relation with each other. In other words, while explicitly trained relations will emerge between B–A and C–A, relations between B–C and C–B will also be derived despite no explicit reinforcement. This ability to respond to arbitrary stimulus relations that are not explicitly derived is a unique feature of human behaviour, and has been found to be a core feature of human language (Lipkens, Hayes & Hayes, 1993; Devaney, Hayes & Nelson, 1986; Barnes, McCullagh & Keenan, 1990).

However, in addition to equivalence relations, a large body of research in Behaviour Analysis has revealed that it is possible for humans to respond in accordance with a variety of different contextually controlled relations, including, for example, Same, Opposite, Different, More than/Less than, and Before/After (e.g., Dymond and Barnes, 1995, 1990; Dymond & Whelan, 2010; Lipkens et al., 1993; Roche & Barnes, 1996). Consequently, a more

appropriate and inclusive account of verbal behaviour has been provided by the behaviour-analytic perspective known as Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001). RFT is a functional contextualist account of human language and cognition, which considers equivalence as just one example of a diverse and extensive range of relations that fall under the phenomenon of Derived Relational Responding (DRR; Hayes et al., 2001; Törneke, 2010). RFT posits that DRR refers to the multitude of ways in which stimuli may be related due to an individual's verbal history of interaction within a social and non-social environment (Leigland, 1997). More specifically, in the same way that reinforcement contingences within a particular history may derive equivalence classes between stimuli (e.g., derived A-C relations), RFT proposes that a there is multitude of ways that different relations may be derived or abstracted between stimuli. Simply put, this essentially refers to the same core process as outlined above (i.e., A-C derived relations), but refers to the variety of different ways in which these relations can be derived (e.g., Same/Opposite, etc.; Hayes, 1994).

1.8 Behaviour Analytic Implicit Measures

In recent years, behaviour-analytic researchers have utilised this paradigm of derived relational responding to develop complex investigative tools for ascertaining verbal histories of relational responding and stimulus relations. The RFT approach to language conceives the social-cognitive 'attitude' concept in terms of a network of trained and derived stimulus relations, established within an individual's verbal and non-verbal interaction histories (Grey & Barnes, 1996; see also Moxon, Keenan, & Hine, 1993; Roche, Barnes-Holmes, Barnes-Holmes, Stewart, & O'Hora, 2002; Schauss, Chase, & Hawkins, 1997). In this way, the measures developed within Behaviour Analysis can be considered analogous to the social-cognitive "implicit attitude" measures, insofar as they are designed to be an indirect means of assessing an individual's histories of relational responding that are both explicitly trained and

derived indirectly within a social and non-social environment. Thus, behaviour-analytic measures work from within a functional and well-defined paradigm, and as such, may circumvent some of the theoretical and conceptual confusion that pertains to the social-cognitive measures (De Houwer, 2011; Hughes et al., 2011).

1.8.1 The Watt, Keenan, Barnes, and Cairns (1991) Stimulus Equivalence Procedure

The first methodology that was developed as a means of assessing histories of relational responding and stimulus relations was the Watt, Keenan, Barnes, and Cairns (1991) Stimulus Equivalence procedure. In brief, the Watt et al. procedure aimed to identify a history of relational responding between Catholic surnames and Protestant symbols among Northern Irish and English individuals. Within this study, participants were exposed to a simple matching-to-sample (MTS) procedure, whereby relations between Catholic family names and nonsense syllables (A-B relations) were trained alongside relations between the same nonsense syllables and Protestant symbols (B-C relations) with the aim of deriving A-B-C equivalence classes. It was expected that due to the tendency among Northern Irish individuals to categorize names on the basis of Protestant or Catholic religions (Cairns, 1984), there would be a resistance observed for the Northern Irish sample to the formation of derived equivalence classes between Catholic names and Protestant symbols relative to the English participants. Indeed, the verbal histories of the these individuals did appear to disrupt the establishment of equivalence relations relative to the English sample, with 12 out of the 19 Northern Irish participants incorrectly pairing a novel Protestant name rather than the expected Catholic name with the Protestant symbols in an equivalence test. This simple equivalence-testing paradigm has been replicated several times, with similar effects observed within a variety of contexts (Barnes, Lawlor, Smeets, & Roche, 1996; Dixon, Rehfeldt, Zlomke, & Robinson, 2006; McGlinchey, Keenan, & Dillenburger, 2000; Merwin & Wilson, 2005; Roche, Ruiz, O'Riordan, & Hand, 2005).

This procedure, therefore, provided clear experimental evidence that an individual's history of relational responding may interfere with the formation of novel relations (O'Reilly et al., 2012). A phenomenon that may explain this finding is the concept of 'Behavioural Momentum' (Nevin & Grace, 2000), a metaphor used in Behaviour Analysis to describe resistance to change effects often observed in behavioural experiments. In essence, the phenomenon of behavioural momentum is considered as analogous to the velocity of a physical body in motion in Newtonian physics, whereby a behaviour that has been consistently reinforced in the learning history of an individual is likely to persist even in spite of a change in environmental conditions (i.e., a change in reinforcement contingencies; Mace, Hock, Lalli, West, Belfiore, Pinter, & Brown, 1988). Thus, behavioural momentum provides an explanation for why there is resistance against behavioural change, such that a disruption to the momentum of a stable or well-established behaviour (e.g., verbal and non-verbal histories of relational responding) is met with greater resistance than behaviours that have not as well established in the history of the individual. In this way, the Watt et al. (1991) study can be conceived as an empirical example of the resistance to change observed when required to form novel equivalence classes (i.e., Catholic names-Protestant symbols) that oppose the existing verbal histories for an individual. Furthermore, the Watt et al. procedure provides considerable evidence that it is indeed possible to develop a measure of an individual's verbal histories of relational responding that does not need to appeal to hypothetical or mentalistic assumptions (i.e., the attitude concept) for explanation.

1.8.2 The Implicit Relational Assessment Procedure

The Watt et al. (1991) methodology provided the first empirical evidence for creating a functionally understood measure capable of ascertaining an individual's history of stimulus relations, thus providing the first viable alternative to traditional explicit measures. More recently, however, the DRR paradigm has been used by behaviour-analytic researchers to

develop more functional, inclusive, and discriminative investigative tools for ascertaining verbal histories of relational responding. The first of these measures is the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes, Hayden, Barnes-Holmes, & Stewart, 2008). The IRAP was developed to expand on the Watt et al. (1991) procedure and to provide a measure capable of more comprehensively assessing derived relations. In brief, the computer-based IRAP methodology requires participants to respond quickly and accurately to a relation between two stimulus categories (e.g., 'Love'/'Hate' and 'Pleasant'/'Unpleasant'), which are presented on-screen in the context of specific relational terms (e.g., 'Similar'/'Opposite'; e.g., Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010). By applying a similar block structure as in the IAT (i.e., where one block is believed to be 'consistent' with the individual's verbal history and another is believed to be 'inconsistent'), the IRAP is proposed to assess an individual's verbal histories of relating these stimuli, resting on the core assumption that participants will respond more quickly to relations that are consistent with their verbal and nonverbal histories (Barnes-Holmes et al., 2010; Roche et al., 2012).

The Relational Elaboration and Coherence (REC) model is a behaviour-analytic interpretation of the behaviours captured by the IRAP, which provides an explanation of how behaviour-analytic theory may be extended to the measurement of an individual's history of relational responding (Barnes-Holmes et al., 2010). Essentially, the REC model posits that the behaviours captured by indirect measures such as the IRAP reflect a relation within the individual's verbal and non-verbal histories between some targeted relation (e.g., Love-Pleasant) and a contextual cue (e.g., the relational term 'Similar'). Within the paradigm of the REC model, therefore, each trial on the IRAP is believed to elicit a brief and immediate relational response that is a function of participants' learning histories. The most probable response is believed to be emitted first and in the shortest time, such that response latencies

should be shorter for trials that are consistent with the individual's learning histories. In contrast, if the trial is inconsistent with their verbal histories (that is, the required response is in opposition with their immediate or most probable response) the REC model posits that response latencies will be lower. Across multiple trials in the IRAP, therefore, the trials that elicit the shortest response latencies are considered consistent with the individual's verbal histories, while in contrast the trials eliciting the longest response latencies are considered inconsistent. In this way, measures such as the IRAP are believed to differ from traditional explicit measures primarily in terms of time pressure (Barnes-Holmes et al., 2010). That is, within the IRAP participants are required to respond quickly and accurately within a particular time frame, and consequently, according to the REC model, the most immediate (i.e., the most probable) responses are emitted. When the time pressure is eliminated, however, participants can engage in longer and more complex relational responding that may be mediated by additional factors, such as socially desirable responding, for example.

However, while the IRAP is making substantial progress in terms of providing a functionally understood implicit test methodology, it has been suggested that one slight limitation of the Watt et al. and IRAP measures is the fact that the procedural implicitness of the test may, in a similar manner to the IAT, be compromised by presenting the specific relation under analysis on-screen during the procedure (Roche et al., 2012). That is, while the exact nature of the test may not be immediately discernible to participants, it is unclear whether or not having the target stimuli on-screen alongside the relational term alerts the participant to the issue under investigation. While one study has directly investigated the fake-ability of the IRAP (i.e., the degree to which the IRAP effect can be diminished by conscious control) and revealed no significant faking effects (McKenna, Barnes-Holmes, Barnes-Holmes, & Stewart, 2007), an additional measure adopting a different approach would undoubtedly contribute to the growing literature on behaviour analytic implicit

measures. Additionally, it could also be argued that the length of time required to complete a typical IRAP, which be up to 60 minutes (e.g., Barnes-Holmes, Murphy, Barnes-Holmes, & Stewart, 2010) may not lend itself as easily to widespread application within practical settings and, consequently, for practical purposes, the development of a shorter measure may be pragmatically beneficial. Lastly, as previously discussed, some researchers have suggested that relying on response accuracy or fluency rather than reaction time (as in the IAT and IRAP methodologies) may provide a more accurate representation of participant difficulty across blocks (Gavin et al., 2012).

1.8.3 The Function Acquisition Speed Test

The second behaviour-analytic measure developed in recent years, which may address these concerns, is a procedure known as the Function Acquisition Speed Test (FAST; O'Reilly, Roche, Ruiz, Ryan, & Campion, in press, O'Reilly et al., 2012). Essentially, the FAST adheres to the same basic core processes of the Watt et al. and IRAP measures, in that it attempts to identify verbal histories of relational responding. However, it works from within a slightly different paradigm than the previous two behaviour-analytic methodologies by attempting to establish functional response classes (i.e., a class based on the functional relations between its responses and classes of antecedent and consequent environmental events) rather than derived relations (O'Reilly et al., 2012). In brief, the FAST establishes a functional response class containing two target stimulus classes by requiring participants to learn, via corrective feedback, a common key-press response (e.g., a left or right key-press) when presented with individual stimuli on-screen (O'Reilly et al., 2012). No instructions are given to the participant prior to the procedure, and the establishment of a function equivalence class is achieved exclusively through reinforcement (i.e., verbal feedback presented on-screen). A similar block structure to the IAT is utilised, wherein each block is comprised of four stimulus classes, which are presented on an individual basis across

multiple trials. In this way, a functional response class may be formed the between two stimulus classes requiring a left key-press response and between the two requiring a right key-press response.

More specifically, as outlined by O'Reilly and colleagues, the general FAST procedure is comprised of four phases. Phase 1 involves simple matching-to-sample training, wherein A1-B1 and A2-B2 equivalence relations between particular stimuli (e.g., two classes of nonsense syllables) are established (O'Reilly et al., 2012). Phases 2, 3, and 4 then involve a short (approximately five to ten minute) testing procedure wherein the strength of the A1 and B1 relations are determined by investigating the rate at which a functional A1-B1 response class can be established relative to non-trained stimulus classes. Phases 2 and 4, therefore, consist of baseline phases that measure the number of trials required to establish response classes between two novel and previously unrelated stimulus classes (e.g., two classes of nonsense syllables between which no relations have been trained). Phase 3 then determines the comparative strength of the trained A1-B1 relation through two separate blocks; one that is consistent with the relations established within the individual's verbal history (i.e., the derived A-B relation) and one that is inconsistent. Specifically the consistent block of Phase 3 investigates the number of trials required to establish two functional response classes between the two previously related classes of nonsense syllables (i.e., A1 and B1 share a common key-press response function), and between two novel and previously unrelated stimulus classes (i.e., N1 and N2 share a common key-press response function). The inconsistent block then investigates the acquisition rates for establishing a response class between stimulus classes that were not previously trained (i.e., A1-N1 and B1-N2 response classes).

The FAST is then assessed entirely based on the comparative acquisition rates (i.e., the number of trials required to establish functional response classes) between the

inconsistent and consistent blocks, which are compared to each other and to the baseline rates to provide an index of pre-existing relational strength (O'Reilly et al., 2012). That is, within the FAST, the rate at which participants acquire a particular level of fluency (i.e., a particular level of accuracy under certain time constraints) across different phases is utilised as the sole criterion for assessing the existence and strength of a particular stimulus relation. In this way, the FAST builds upon both the aforementioned resistance to change effects (Nevin & Grace, 2000) and the finding that pre-existing equivalence classes can interfere with the formation of new equivalence classes that oppose them (Watt et al., 1991) to assess an individual's verbal histories, such that the degree of resistance that is met when establishing a functional response class containing a particular stimulus relation may reflect the extent to which that relation contradicts the individual's verbal histories.

The FAST emerged organically from research investigating how to functionally analyse the IAT effect within an explicit behaviour-analytic perspective (e.g., Gavin, Roche, & Ruiz, 2008; Gavin, Roche, Ruiz, Hogan, & O'Reilly, 2012; O'Toole & Barnes-Holmes, 2007; Ridgeway, Roche, Gavin, & Ruiz, 2010; Roche et al., 2005; see also Hall, Mitchell, Graham, & Laves, 2003; Mitchell, Anderson, & Lovebird, 2003 for associative conditioning approaches) and, as such, may provide a quick, viable, and functionally understood alternative to existing measures such as the IAT. Firstly, unlike the IAT, the FAST was developed from within a functionally understood paradigm wherein its underlying assumptions are entirely testable (i.e., stimulus equivalence and derived relational responding; Gavin et al., 2012; O'Reilly et al., 2012). That is, earlier research that led to the development of the FAST suggested that it was indeed possible to measure both trained and derived stimulus-stimulus relations using IAT-type procedures (Gavin et al., 2008). According to Gavin et al. (2008), the IAT may in fact function as a subtle test for derived relations, rather than as a measure of unconscious mental associations *per se*. Thus, the

FAST, in a similar manner to the IRAP and Watt et al., procedures, provides empirical support for the idea that it is entirely possible to develop a measure based on testable and functionally understood assumptions rather than appealing to mentalistic or hypothetical explanations (O'Reilly et al., in press). Secondly, performance on the FAST is assessed based on trial requirements and participant fluency (combined accuracy and time), and, as such, does not rely on statistical manipulations to distort the data, nor does it utilise a complex and somewhat opaque scoring algorithm (O'Reilly et al., 2012). Lastly, unlike the IAT, the FAST is capable of ascertaining the pre-existing strength of a single stimulus relation in isolation, thus providing an absolute rather than relative assessment of verbal histories (O'Reilly et al., 2012).

The FAST also addresses some of the issues outlined above with the Watt et al. and IRAP methodologies in that the procedure establishes functional response classes entirely through reinforcement and thus may be more procedurally implicit than if the two stimuli and a relational term were presented on-screen. Additionally, the FAST is considerably less demanding of participants in terms of both attention (there are no relations to derive) and time (without MTS training the FAST takes approximately 5-10 minutes to administer, relative to 60) than the IRAP procedure, though it should be noted that the aim of the IRAP is to provide a comprehensive account of a verbal stimulus relations, and as such it is designed to extrapolate the multitude of complex stimulus relations that may exist between two sets of stimuli. In contrast, the FAST is designed as a quick and easy-to-administer assessment of the existence and strength of a specific stimulus relation and thus may not provide as substantive an analysis as the IRAP. However, the relative speed of the FAST does ensure its practical utility within a wide range of contexts and, as such, its development may address a gap within the behaviour-analytic research literature.

1.9 Developing the Function Acquisition Speed Test for Homonegativity

The current thesis aims to move the Function Acquisition Speed Test (FAST) into the context of pre-existing real world stimulus relations. Specifically, this research aims to develop the FAST procedure within the arena of sexual prejudice as a means to assess the existence and strength of anti-gay verbal histories. In other words, this research is concerned with the development of a measure of anti-gay implicit attitudes, with attitudes conceived here in terms of verbal relations between stimuli, and the transformation of the functions of those stimuli in accordance with the network (Roche et al., 2012). For convenience purposes, the term attitudes will henceforth be used interchangeably with this description for the remainder of this thesis.

1.9.1 Homonegative Prejudice

Psychological research into the study and measurement of anti-gay attitudes is a relatively recent phenomenon. The concept of 'homophobia' was first proposed by Weinberg (1972) to define a fear or hatred of homosexual individuals. In the years that have followed, however, the concept of homophobia has evolved in the literature to signify any negative belief, action, attitude, or discrimination toward Lesbian, Gay or Bisexual (LGB) individuals (Plugge-Foust & Strickland 2000). Thus, a number of researchers have moved away from using the term 'homophobia', which is indicative of a phobia or fear, preferring instead to utilise terms that more accurately reflect anti-gay attitudes or prejudice (e.g. Logan, 1996; Ahmad & Bhugra, 2010). Terms such as heterosexism, sexual prejudice, sexual stigma and homoprejudice have all been used frequently in the literature to refer to anti-gay bias, although the most commonly used term is 'homonegativity', which Morrison and Morrison (2002) propose encompasses all prejudicial, derogatory, or negative beliefs or attitudes toward LGB individuals.

The issue of homonegative prejudice against LGB individuals is a considerably pressing social problem. Firstly, the International Lesbian and Gay Alliance (ILGA) revealed

in its most recent 2013 report that Homosexuality remains illegal in 76 countries across the globe, with Homosexual acts punishable by the death penalty in 12 of those 76 (Itaborahy & Zhu, 2013). Furthermore, as of May 2013, this report showed that the number of countries currently extending full civil and legal marriage equality for LGB individuals is just 14, with only 12 of these allowing legal adoption for same-sex couples. Additionally, according to this report, just 6 countries worldwide constitutionally prohibit civil discrimination on the grounds of sexual orientation.

International statistics concerning the mental and physical health of LGB individuals further point to the severity of this issue. One study, for example, found a strongly significant increased risk of premature death among LGB individuals, estimating that, among the Canadian LGB community, there was an estimated 818-968 deaths per year as a result of suicide, 1232-2599 deaths per year as a result of smoking; 236-1843 deaths per year as a result of alcohol abuse; and 64-74 deaths per year as a result of illicit drug use (Banks, 2003). Furthermore, several studies have revealed that LGB individuals typically report significantly higher levels of sexual, verbal, and physical abuse as a result of their LGB orientation (Balsam, Beauchaine, Mickey, & Rothblum, 2005; Berrill, 1992; D'Augelli, Grossman, & Starks, 2006). Additionally, research suggests that, when compared with their heterosexual counterparts, Gay men and Lesbian women report higher instances of mental health problems, including substance use disorders, affective disorders, and suicide (e.g., Cochran, 2001; Sandfort, de Graaf, Bijl, & Schnabel, 2001). It has also been found that LGB adolescents who have experienced homonegative bullying or victimization as a result of their LGB identification are at an increased risk of a wide range of mental and physical health problems (Coker, Austin, & Schuster, 2010). Similarly, research suggests that the internalization of societal homonegative beliefs (i.e., 'internalized homonegativity') among LGB individuals that results from persistent discrimination may exert significant negative

impact on their overall mental health and well being (Allen & Oleson, 1999; Herek, Cogan, Gillis, & Glunt, 1998; Meyer & Dean, 1998; Rowen & Malcolm, 2002). Several researchers have posited that the increased instance of physical and mental health problems among LGB populations is largely the result of the stigma, prejudice, and discrimination experienced by LGB individuals, and the associated stress of existing within a stigmatized minority group (Friedman, 1999; Meyer, 2003).

Additionally, within an Irish context specifically, the issue of LGB discrimination remains a prevalent and problematic social issue. Although there has been a growing effort in recent years to promote acceptance of homosexuality in Ireland, homonegativity remains a persistent social problem. Firstly, Ireland does not currently extend full marriage and adoption equality to same-sex couples, nor it is unconstitutional to discriminate against an LGB individual on the grounds of sexual orientation (Itaborahy & Zhu, 2013). Furthermore, a relatively recent study by Maycock, Bryan, Carr, and Kitching (2009) surveyed 1100 LGBT individuals in Ireland, aged between 14-73, and revealed that, as a result of their LGBT identification, 80% of survey participants reported experiencing verbal abuse, and 25% of survey participants reported experiencing physical abuse over their lifetime. 40% of survey respondents reported having received physical threats because they were, or thought to be, LGBT. Just under 8% of respondents reported being attacked with a weapon or implement on at least one occasion, while 9% admitted being attacked sexually as a consequence of their LGBT identification. The study also highlighted the increased risk of suicidality in those who have experienced homophobic bullying and other forms of victimization as a result of their LGBT identification, with 17.7% of survey respondents and almost a third of 40 in-depth interview participants admitted attempting suicide at least once in their lifetime. Statistically significant associations were also found between lifetime suicide ideation and

having been the victim of homophobic bullying or attacks as a result of their LGBT identification.

Research also indicates that overt anti-gay bullying and victimization are significant social problems in Irish society, particularly in second-level schools. For example, Norman (2005) reviewed the issue of homonegativity in second-level Irish schools through the experiences of teachers. The study found that 79% of teachers surveyed reported encountering homophobic verbal bullying in schools, while 30% of these had encountered this type of bullying on more than ten occasions in the previous term. Similarly, Minton, Dahl, O' Moore, and Tuck (2008) found that in a survey of LGB youth in Irish second-level schools, half of all respondents reported instances of homophobic bullying during the previous three months, while 30% reported bullying during the 5 days previous. An additional study conducted by O'Higgins-Norman, Galvin, and McNamara (2006) investigated the use of words such as 'faggot', 'queer', 'Gay' and 'dike' as a means to insult each other in Irish schools, and found that all 127 students interviewed described the frequent use of such terms. Similarly, Kehily (2001), revealed high instances of homonegative language among students, with homonegative jokes, 'banter', and abuse featuring regularly within male peer groups, noting "the potentially emasculating experience of being called Gay" in second-level Irish schools (p. 121). In addition to the problem of overt homonegative behaviour, some researchers have examined the more subtle or covert forms of homonegativity in Ireland. Mac an Ghaill, Hanafin and Conway (2004), for instance, noted the casual but frequent use of homonegative pejoratives in Irish society, which Minton and colleagues (2008) argue reflects the more insidious and non-targeted general undercurrent of homophobia that pervades Irish workplaces and schools.

1.9.2 The Measurement of Homonegativity

Traditionally within Psychology, as was the case with all attitude research, an individual's beliefs toward homosexuality were assessed using standard self-report measures such as questionnaires or surveys (De Houwer, 2006). The most popular self-report measures that have been developed in recent decades for homonegative attitudes include the Hudson and Ricketts (1980) Index of Homophobia, the Attitudes toward Gay men and Lesbian women scale (ATLG, Herek, 1988, 1994), which has been particularly utilised widely among attitude researchers since its development (e.g., Hegarty, 2002; Nierman, Thompson, Bryan, & Mahaffey, 2007; Mohipp & Morry, 2004; Schellenberg, Hirt, & Sears, 1999; Simon, 1995; Simoni, 1996), and the relatively recent Modern Homonegativity Scale (MHS; Morrison & Morrison, 2002), which was developed as a means to provide a more subtle and indirect measure of homonegativity than previous scales. That is, while the ATLG contains terms such as *"Male homosexuality is a perversion"*, the MHS utilises items such a *"Gay men should stop shoving their lifestyle down other people's throats"* in order to identify more contemporary homonegative attitudes that may not be as strongly governed by moral or traditional objections to homosexuality (Morrison, Kenny, & Harrington, 2005).

Interestingly, the explicit attitudinal research that has been conducted over the past three decades has revealed a substantial reduction in Western societies in self-reported levels of homonegativity (see Cullen & Barnes-Holmes, 2008; Herek, 2000; Hicks and Lee, 2006; Steffens and Wagner, 2004; Yang, 1997, for reviews). However, based on the wealth of literature on LGB discrimination outlined above, it would seem somewhat paradoxical to interpret these findings as indicative of the amelioration of homonegativity in the West. When the aforementioned problems inherent in self-report measures (i.e., social desirability) are taken into account, however, the above findings may be understood in terms of a reduction in *overt* homonegative attitudes, rather than homonegativity *per se*. Consequently, it is necessary to identify a methodology that can circumvent this issue and thus potentially

provide a more accurate reflection of modern anti-gay attitudes. While more subtle measures such as the MHS have attempted to improve upon this issue, the fact that the measure still relies on self-reported levels of homonegativity suggests it may still fall prey to the issues plaguing explicit measures generally. Furthermore, it has been argued that even when more indirect measures (e.g., the MHS) are used, it may still be possible to respond in a socially desirable manner when one is aware that subtle prejudicial attitudes are under investigation (Fazio, 1995). Thus, in recent years a small body of research within both Behaviour Analysis and Implicit Cognition has been devoted to the development of an alternative indirect measure of homonegativity (Cullen & Barnes-Holmes, 2008).

This research has predominantly been conducted within the field of Social Cognition using the IAT, with a small number of published studies devoted to the development of an implicit attitudes of homonegativity (Banse et al., 2001; Hudepohl, Parrott, & Zeichner, 2010; Jellison, McConnell, & Gabriel, 2004; Jost, Banaji, & Nosek, 2004; Nosek et al., 2005, 2007; Rohner & Björklund, 2006; Rowatt, Tsang, Kelly, LaMartina, McCullers, & McKinley, 2006; Steffens, 2004; Steffens & Buchner, 2003), in addition to one unpublished IRAP measure of homonegativity (Cullen, Barnes-Holmes, Barnes-Holmes, and Stewart, 2007; in, Cullen & Barnes-Holmes, 2008).

However, in light of the previously discussed concerns regarding the assumptions, presentation format and scoring method of the IAT, a functionally understood implicit measure is undoubtedly required. Additionally, though the IRAP has made significant progress in this regard in that it has provided a behaviour-analytic implicit measure of homonegative attitudes, the previous IRAP research in this area has been entirely preference-based (i.e., Gay-Straight preference), as has the IAT (see Cullen & Barnes-Holmes, 2008). While the IRAP does enable a breakdown of individual blocks to provide a more absolute index of relational strength (see Barnes-Holmes et al., 2010; Drake et al., 2010), to date, no

Homonegativity IRAP has been developed outside of a relativistic paradigm. Thus, it would be interesting to develop a functionally understood test capable of investigating homonegative verbal histories in isolation. Furthermore, given the potential practical utility of a measure that can quickly and reliably assess the strength of pre-existing homonegative verbal histories, it would be pragmatically beneficial to develop a measure that is both fast and easy-to-administer to large populations.

1.9.3 Current Research

Consequently, the current research aims to add to this small body of research and develop the Function Acquisition Speed Test within the context of homonegative prejudice. That is, this research aims to develop the FAST as a measure of existing homonegative stimulus-stimulus relations. Thus, this research will adhere to the typical FAST procedure outlined previously, with the only exception that it will test for existing relations that may have been derived within an individual's social and non-social histories, rather than those that have been established via MTS training. In brief, this will be achieved by establishing distinct response functions for stimuli expected of being related as a result of prior social interaction. The individual's homonegative verbal histories will then be assessed using the simple paradigm outlined above, wherein the degree of behavioural resistance to the formation of a functional Gay-Good stimulus class should reflect pre-existing levels of homonegativity, such that the Gay-Good relation should be orthogonal to the existing verbal histories for strongly homonegative individuals. In this way, through the development of a 'Homopositivity FAST', this thesis proposes to develop a functionally understood, quick, and effective means of assessing homonegative prejudice.

Chapter 2

Identifying Exemplars for a Homopositivity FAST: Comparing the Relative Effectiveness of Verbal versus Pictorial Stimuli

Experiments 1 and 2

2.1 Experiment 1

2.1.1 Introduction

The aim of the current research is to examine the degree of behavioural resistance to the formation of functional Gay-Good stimulus classes (e.g., 'Gay people' and 'positive attributes'). When developing the FAST within this specific context of homonegativity, one of the first factors that must therefore be addressed is the experimental stimuli that will be employed. The general purpose of implicit measures, according to the social-cognitive and behaviour-analytic literature, is to assess the strength of relation (or association) between two concepts, as represented by some experimental stimuli (e.g., Bluemke & Friese, 2005; Foroni and Bel-Bahar, 2010). That is, an assumption in the implicit testing literature is that these measures are assessing the strength of the relation between the general stimulus categories (e.g., Gay people and positive evaluative terms) rather than merely the individual associations between specific exemplars (e.g., this *particular* image of a Gay person and the word 'good'). Thus, implicit testing research rests on the general principle that observed effects are modulated by how well the stimuli employed represent the categories of interest (e.g., De Houwer, 2002; Govan & Williams, 2004). Indeed, as outlined by O'Reilly and colleagues (2012), the reliability of any FAST developed within a real-world context will be largely contingent upon the representativeness of the stimuli chosen for the target categories. Therefore, for the Homopositivity FAST to be conceptually valid and reliable, it is crucial to ensure that the stimuli selected are optimally representative of the nominal categories according to which participants will categorise target stimuli (i.e., the images or words representing the concepts 'Gay and 'Good'; Gawronski & Payne, 2010).

The importance of selecting appropriate exemplars to represent the attributes and categories of direct interest has been widely addressed in the implicit testing literature (e.g., Govan & Williams, 2004; Mitchell, Nosek, & Banaji, 2003; Steffens & Plewe, 2001),

particularly given the noted modulatory effect of stimuli on test outcomes (e.g., Foroni & Bel-Bahar, 2010; Gawronski & Payne, 2010). As previously discussed, research has suggested that the nature of the stimuli employed in implicit test methodologies may exert significant influence on both the direction and magnitude of IAT effects (e.g., Bluemke & Friese, 2005; De Houwer, 2002; Govan & Williams, 2004; Steffens & Plewe, 2001). Furthermore, it has been found that there have been discrepancies in outcomes with different versions of the same IAT (e.g., a race-IAT) depending on the chosen exemplars (Banaji, 2001; Steffens & Buchner, 2003; Lane et al., 2007; Foroni & Bel-Bahar, 2010). Consequently, some researchers have suggested that variances in levels of representativeness of stimuli may partly determine IAT outcomes, such that different stimuli may elicit different representations quite distinct from that which was intended (Blair, 2002; Devos & Ma, 2008, see also De Houwer, 2002; Govan & Williams, 2004). However, as argued by Bluemke and Friese (2005), despite there being no uniform set of rules for the selection of exemplars, different versions of the same IAT are often viewed as interchangeable. The current research, therefore, aims to experimentally deduce the most appropriate and representative exemplars for a measure examining attitudes toward homosexual individuals.

One of the first factors that will be investigated here is the relative capability of pictorial versus verbal stimuli to represent the concepts ‘Gay’ and ‘Good’ in a Homopositivity FAST. Previous research has suggested that an important distinction exists between stimulus exemplars that contain what is referred to as high and low levels of representation (Foroni & Bel-Bahar, 2010). Stimuli with low levels of representation are proposed to require lower levels of processing relative to more highly representative stimuli, which require more perceptual and semantic processing (e.g., Nelson & Schreiber, 1992; Paivio, Walsh, & Bons, 1994; Schwanenflugel & Shoben, 1983). Researchers posit that semantic-verbal stimuli fall into the category of stimuli with high levels of representation, given that they are proposed to

further generate additional related exemplars and other visual features associated with the category, while more specific perceptual-visual stimuli are considered to have comparatively low levels of representation (Marsolek, 1999; Tanaka, Luu, Weisbrod, & Kiefer, 1999; Whatmough, Verret, Fung, & Chertkow, 2004). It has been suggested, for example, that to process a single forename (e.g., 'John') may generate further information and knowledge related to that particular subgroup or category (e.g., 'information about people you know named John'), relative to knowledge generated by a single face stimulus (Park, Ryan, & Judd, 1992; Rothbart & John, 1985). Research has supported this assumption, suggesting that perceptual pictorial stimuli and semantic word stimuli may actually be processed and engaged with differently (Henson, Price, Rugg, Turner, & Friston, 2002; Ito & Ulrand, 2005; Rossion, Joyce, Cottrell, & Tarr, 2003; Todorov, Gobbini, Evans, & Haxby, 2007).

While the research is somewhat conflicted on the relative effectiveness of pictorial and verbal stimuli within implicit tests (see Cullen & Barnes-Holmes, 2008), IAT research appears to support the foregoing hypothesis, typically reporting higher effect sizes for verbal or word-IATs than for pictorial IATs (see Mitchell, Nosek, & Banaji, 2003; Hurtado, Haye, González, Manes, & Ibáñez, 2009; Nosek, Banaji, & Greenwald, 2002). Foroni and Bel-Bahar (2010) attribute this finding to the above distinction between stimuli with high levels of representation and those with low levels, suggesting that verbal stimuli (e.g., a forename associated with Black-American individuals) generate knowledge related to the wider stimulus category and, as such, are more generally representative of the overarching nominal category (Black people) than a pictorial image of a specific Black individual. As a starting point, therefore, in the identification of the most appropriate stimulus set for the development of a homonegativity FAST, the current research aims to build on the above findings and first establish the appropriateness of unambiguous verbal stimuli related to the target stimulus categories of interest.

The next important factor to consider when deducing the most suitable set of stimuli for a Homopositivity FAST is the number of exemplars that will represent the categories of interest. While little research has examined this directly, a small body of IAT research may be useful in guiding the selection of exemplars in the current experiment. Nosek, Greenwald and Banaji (2005), for instance, note that the specific number of exemplars chosen for an implicit measure is usually arbitrary, suggesting that it is usually determined by how easily exemplars can be identified that adequately represent the target category. Indeed, previous IAT research has suggested that the number of exemplars per stimulus category does not significantly influence the magnitude of the IAT effect, such that effects appear to be identical whether 5 or 25 exemplars are employed (Greenwald et al., 1998; Nosek, 2005). However, research has suggested that the IAT effect is somewhat reduced when very few exemplars are utilised. Nosek and colleagues (2005), for example, found that effects across multiple IATs appeared to be minimized when exemplars were reduced to just one or two per category. Consequently, these researchers suggest that using multiple rather than single exemplars may reduce ambiguity and ensure a more accurate representation of the target category (Nosek et al., 2005). Additionally, these researchers note the importance of ensuring that each exemplar used is an excellent and appropriate representation of the stimulus category, and that the representation of the super ordinate category is not diluted by items that do not adequately depict the stimulus category in the desired way (Nosek et al., 2005). When dealing with a complex issue such as social prejudice, the identification of multiple target stimuli that are each highly representative would be somewhat challenging. Consequently, rather than using a large number of potentially weak exemplars, the current research will utilise four semantically representative exemplars per stimulus class, thus ensuring that each stimulus is optimally representative of the category while also minimising any potential ambiguity surrounding the evaluation of the stimulus classes.

In order to assess the effectiveness of the exemplars chosen to represent the stimuli categories, the current experiment will adopt what is known as a "known-groups" approach (Nettler & Golding, 1946), a validation technique commonly employed in attitude research (e.g., Greenwald et al., 1998; Pettegrew & Wolf, 1981; Petróczi, Aidman & Nepusz, 2008; Rest, 1977). Essentially, this refers to the use of a sample in which the participants' beliefs or affective responses to some variable of interest are known (e.g., two groups differing in their self-reported beliefs toward the elderly), and investigating whether or not the measure can successfully discriminate between the two groups (Cronbach & Meehl, 1955). A considerable proportion of IAT research has adopted this approach to validity, focusing on whether or not IAT scores between two distinct groups of individuals ultimately differ in the expected way (Fazio & Olsen, 2003). For example, this approach has been employed when validating IATs concerned with phobias, by comparing scores between phobic versus non-phobic individuals (Teachmann, Gregg & Woody, 2001), and similarly for sexual orientation IATs, between self-identified Gay and straight individuals (Banse et al., 2001). In this way, the suitability of the experimental stimuli can be ascertained, such that FAST scores from those with high self-reported homonegativity should theoretically differ from the rest of the sample.

It is important to note that while the known groups method may be suitable for certain types of research (e.g., assessing difference between phobic versus non-phobic individuals), it suffers from limitations in he context of socially sensitive attitudes, such as anti-gay prejudice. As previously discussed, traditional explicit measures are typically prone to self-presentational distortions, due to a tendency for people to respond in a socially desirable manner (e.g., Greenwald et al., 1998). As such, relying on a questionnaire or survey to conclusively establish a group of individuals whose degree of anti-gay prejudice is definitively "known" would be somewhat misguided. However, it would appear that when

these distortions occur, they are typically in the socially desirable direction (i.e., with a tendency to report lower levels of prejudice than may be warranted by corresponding attitudes) rather than leading to a general distortion of results (e.g., Paunonen & Le Bel, 2012). Thus, while self-report measures may then be an inaccurate reflection of attitudes for those who report low levels of homonegativity, it may be useful in identifying strongly anti-gay individuals, given that their responses may be somewhat less prone to social desirability biases. Additionally, the current research aims to assess explicit levels of so-called 'modern' homonegativity, which has not been found to be associated with social desirability to the same extent as more traditional forms of anti-gay attitudes (Morrison & Morrison, 2002).

To summarise, the current experiment aims to assess the effectiveness of multi-exemplar verbal stimulus categories in a FAST for homopositive bias. The effectiveness of these stimuli will be assessed largely through the use of a known-groups' paradigm, which will involve distinguishing between participants on the basis of self-reported homonegativity and investigating whether FAST scores between these two groups differ in the expected direction. Therefore, it is expected that there may be general patterns visible within the data based on explicit levels of homonegativity, which will be assessed using a pencil-and-paper measure. More specifically, it is expected that there will be significant resistance in the formation of a functional Gay-Good verbal class for strongly anti-gay individuals, given that it should be inconsistent with their verbal histories. Consequently, it is expected that for those with moderate-to-high levels of explicit homonegativity, more trials will be required for the consistent (i.e., the pro-gay) block relative to the inconsistent block. In contrast, a more varied outcome in terms of trial requirements is expected for those with lower self-report levels of anti-gay attitudes, largely due to the self-presentational distortions commonly encountered with explicit scales (discussed previously). Finally, the current experiment aims to assess general levels of anti-gay bias in the sample as a whole, as detected by the current

version of the Homopositivity FAST. As discussed in detail in the introduction chapter, prejudice against LGB individuals would still appear to be a pervasive social problem in Ireland (e.g., Minton et al., 2008; O'Higgins-Norman et al., 2006) and, therefore, it is expected that there will be an overall trend toward generally weak pre-existing Gay-Good stimulus relations for the current sample.

2.1.2 Method

2.1.2.1 Participants

Participants for this experiment were recruited from either the NUI, Maynooth undergraduate population or from acquaintances of the experimenter via word of mouth recruitment. All participants were adults (over 18 years of age) who did not suffer from any condition that would interfere with their ability to perform routine cognitive tasks (this is with the exception of standard age-related cognitive decline in older participants). The sample (n = 40) was divided into two groups on the basis of self-reported homonegativity (i.e., around the median score on the Modern Homonegativity Scale; see Procedure), with participants assigned non-randomly to either Group One (low levels of explicit homonegativity) or Group Two (moderate-to-high levels of homonegativity). Group One (n = 20) was comprised of 10 males and 10 females aged between 20 and 67 (M = 38.55; SD = 16.42), while Group Two (n = 20) was comprised of 8 males and 12 females aged between 21 and 75 (M = 52.55; SD=20.1). [1] The current research also imposed an exclusion criterion for those participants who failed to reach fluency within 200 trials on the first baseline block (Phase 2). As a result, two participants were excluded from the data set. A third participant withdrew their participation during the first baseline phase and consequently this data was also omitted from the experiment.

---

1 Katie Lattimore, an undergraduate student at the National University of Ireland, Maynooth, assisted with data collection as part of her final year research project

2.1.2.2 Ethical Considerations

All elements of this research process received prior approval from the NUI, Maynooth Ethical Committee, and were all in accordance with the Psychological Society of Ireland (PSI) code of ethics, and with the NUI, Maynooth Department of Psychology guidelines for safe working practice in psychological research. However, it should be briefly noted here that the experiments in this thesis do not pose any ethical concerns, given that they were in no way emotionally or cognitively challenging for participants, nor did they permit the diagnosis of any condition or risk. All tasks within this experiment were standard cognitive tasks, and did not allow for the gathering of any sensitive information about individuals. This is assured in the current research given that all information was analysed at a group level (i.e., participants' data was anonymised). The only possible risk to participants in this research was the possibility of minor embarrassment due to the slightly sexual nature of the experimental stimuli and the self-report questionnaire (discussed below), although such concerns are commonplace in all research concerned with attitudes toward sexual matters.

2.1.2.3 Materials

2.1.2.3.1 The FAST Software Package and Stimulus Set

The version of the FAST utilised in the current experiment was developed with the computer programming software Livecode™, which is a multi-platform coding application, and was presented on a 15" laptop (1920×1080 pixel resolution). The FAST employed verbal stimuli as exemplars of the categories of interest, with four exemplars per stimulus category. All stimuli were matched for word length and complexity across each block in order to ensure that certain words did not form a natural response class outside the laboratory. Additionally, across each block, stimuli were excluded if they formed a natural class with additional stimuli from other categories based on colloquial associations. For example, in the target FAST block, the colour blue and the days Monday and Friday were excluded due to

colloquial connotations between the these days of the week and particular positive or negative moods or feelings (e.g. 'feeling blue'/ 'Friday feeling', etc.). With the exception of the target categories ('Gay' and 'Good'), all words employed in this experiment were common, everyday words that were deemed neutral in terms of their evaluative functions. Stimuli for the target categories were chosen carefully based on their perceived semantic representativeness of the categories of interest. For the stimuli representing the category 'Gay', descriptive words were carefully selected from those used within the LGB community that were non-intentionally offensive while still being representative of the category. For the 'Good' category, unambiguously positive and emotive 'feeling' words were selected based on the list of positive words used by Greenwald et al., (1998).

Stimuli for each phase of the FAST, their alphanumeric category code and their category descriptions are listed below (Table 2.1.1). In this research, A1 and B1 refer to the target stimulus categories used (i.e., 'Gay' and 'Good'), while N1 and N2 refer to the two additional categories used during test blocks. Additionally, P and Q refer to the categories used during the practice blocks, while X and Y refer to the baseline stimulus categories.

Table 2.1.1: Verbal exemplars employed in Experiment 1

Alphanumeric	**Stimulus Category**	**Category Description**
Phase 1: Practice Block		
P1	Common, everyday fruits	Lime, Pear, Peach, Plum
P2	Recognisable body parts	Arm, Leg, Hand, Foot
Q1	Recognisable animals	Cow, Pig, Horse, Sheep
Q2	Everyday items of clothing	Pants, Dress, Shirt, Coat
Phase 2: Baseline 1		
X1	Recognisable Western European countries	England, Finland, Sweden, Poland

X2	Common types of birds	Cuckoo, Magpie, Bluebird, Sparrow
Y1	Recognisable sports	Football, Tennis, Baseball, Hockey
Y2	Well known occupations	Baker, Butcher, Teacher, Doctor
Phase 3: Test Block		
A1	Descriptive words for Gay people	Gay, Homo, Homosexual, Queer
B1	Positive emotions/feelings	Love, Happiness, Pleasure, Joy
N1	Common, frequently encountered colours	Red, Yellow, Green, Purple
N2	Days of the week	Saturday, Wednesday, Tuesday, Thursday
Phase 4: Baseline 2		
X3	Common, frequently encountered types of trees	Elm, Birch, Beech, Oak
X4	Everyday kitchen dishes	Bowl, Plate, Cup, Jug
Y3	Common types of fish	Sole, Plaice, Cod, Trout
Y4	Single-digit numbers one to four	One, Two, Three, Four

2.1.2.3.2 The Modern Homonegativity Scale

The current experiment aimed to compare the scores on the Homopositivity FAST between two groups who were divided on the basis of their self-reported levels of homonegativity. Therefore, it was necessary to utilise a measure that could discern the participant's level of self-reported, or explicit, homonegativity. Given that it was the most indirect and subtle measure available, in addition to its reported high levels of reliability and internal consistency, the previously discussed Modern Homonegativity Scale (MHS; Morrison & Morrison, 2002; Morrison et al., 2005) was utilised for this purpose.

The MHS scale is a questionnaire that is based on Morrison and Morrison's framework of modern prejudice, in which contemporary attitudes to LGB individuals are

conceptualised in terms of two distinct forms: old-fashioned and modern homonegativity. The first form of homonegativity is proposed to reflect more traditional moral or religious objections toward Gay individuals. This is then proposed to be in contrast to more modern prejudicial assumptions that surround LGB populations, including, for example, the beliefs that Gay men and women make unnecessary demands on the status quo, that they "advertise" their homosexuality and as such should be considered responsible for their exclusion from mainstream society, and that the issue of sexual discrimination is no longer a pressing social problem. The MHS employed in the current research is a 10-item questionnaire that assesses these levels of modern homonegativity toward Gay individuals, and was developed by Morrison, Kenny, and Harrington (2005; Appendix 2). The questionnaire exists in two parallel forms, with one for attitudes toward Gay men and another toward Lesbian women. However, the literature does suggest a distinction between attitudes toward Gay men and Lesbian women and, as such, it is important that these constructs are examined independently (e.g., Herek, 1984). Specifically, the current experiment utilised the more widely employed attitudes toward Gay men scale (Morrison et al., 2005), though it should of course be noted that this experiment is thusly concerned with Gay men exclusively and not Lesbian women. With regard to scoring, the MHS scale uses a five-point Likert scale (*strongly disagree, disagree, undecided, agree,* and *strongly agree),* with total scale scores ranging from 10 to 50.

2.1.2.4 Design

The dominant research paradigm employed in this research generally was an experimental one, behaviour-analytic in philosophical orientation, and inductive, rather than theory-driven. However, the current experiment employed a traditional between-groups quasi-experimental design with homonegativity scores forming one independent variable and the two critical consistent and inconsistent block types of the FAST together forming a second variable controlling test performance.

2.1.2.5 General Experimental Sequence

2.1.2.5.1 Informed Consent

Prior to any involvement all participants were informed that they are taking part in an experiment assessing a new measure of attitudes that will require them to fill out a short questionnaire and to complete a computer-based learning task, during which a variety of stimuli (including some that are sexual in nature) will be presented to them on the screen. Participants were informed of the format of the task (i.e., that a specific type of word or image will be presented and that a response will be required in the form of a button press, as instructed by the computer program), and of the likely duration of the experiment. Participants were then required to sign a consent form (Appendix 1), and were fully assured that they were not obliged to participate in the research even after having signed the form.

2.1.2.5.2 Function Acquisition Speed Test

Upon supplying written consent, the participant was then required to complete the FAST portion of the experiment. This involved firstly ensuring the participant was comfortable, the environment was quiet, and that there were limited external distractions. The participant was then seated at a desk and presented with the FAST programme on a laptop computer. Each phase of the FAST commenced with a series of instructions on the screen informing the participants of the general nature of the task (see below). These remained on-screen until participants felt confident they understood the instructions, after which they pressed any key on the keyboard to begin (as instructed). For all blocks but the practice block, once the participants began the block, they were presented with a series of stimuli in a randomised order. Each stimulus was presented for 3 seconds, with all word stimuli presented in black 48-type font in the centre of the screen. The shared response function in each block was reinforced via corrective feedback, which was presented after the participant's key-press response. The feedback ('Correct' or 'Wrong') was presented on-

screen for a period of 1.5 seconds in the centre of the screen in red 48-type font. If no response was emitted from participants within the 3-second response window (i.e., no key was pressed), a time-out response ('Wrong') was presented in the centre of the screen, again in red 48-point font. Each block concluded once the participant reached the fluency criterion (i.e., when they responded correctly for ten consecutive trials). Alternatively, the block ceased once the participant failed to reach fluency within 200 trials, after which the next block immediately commenced. This is with the exception of Phase 1, during which participants were required to complete 16 practice trials only and no fluency criterion was applied.

Phase 1: Practice block

The first phase of the FAST is the practice block, during which participants were familiarised with the procedure and the general format of the task. During the practice block, relations between neutral, previously unrelated stimulus categories were established through reinforced shared response functions. Specifically, the stimulus categories P1 (fruit) and P2 (body parts) shared the response key 'Z', while categories Q1 (animals) and Q2 (items of clothing) shared the response key 'M'. The practice block commenced by presenting the participants with the following set of instructions on the computer screen:

> *In the following section your task is to learn which button to press when a word appears on screen. IMPORTANT: During this phase you should press only the Z or the M key. Please locate them on the keyboard now. This initial session is to give you some practice with the task. To help you learn you will be provided with feedback telling you if you are right or wrong. If you have any questions please ask the researcher now. Press any key when you are ready to begin.*

Once the participant began the block, they were presented with 16 trials, after which this portion of the FAST concluded and the instructions for the next block were presented on-screen, regardless of the participant's performance.

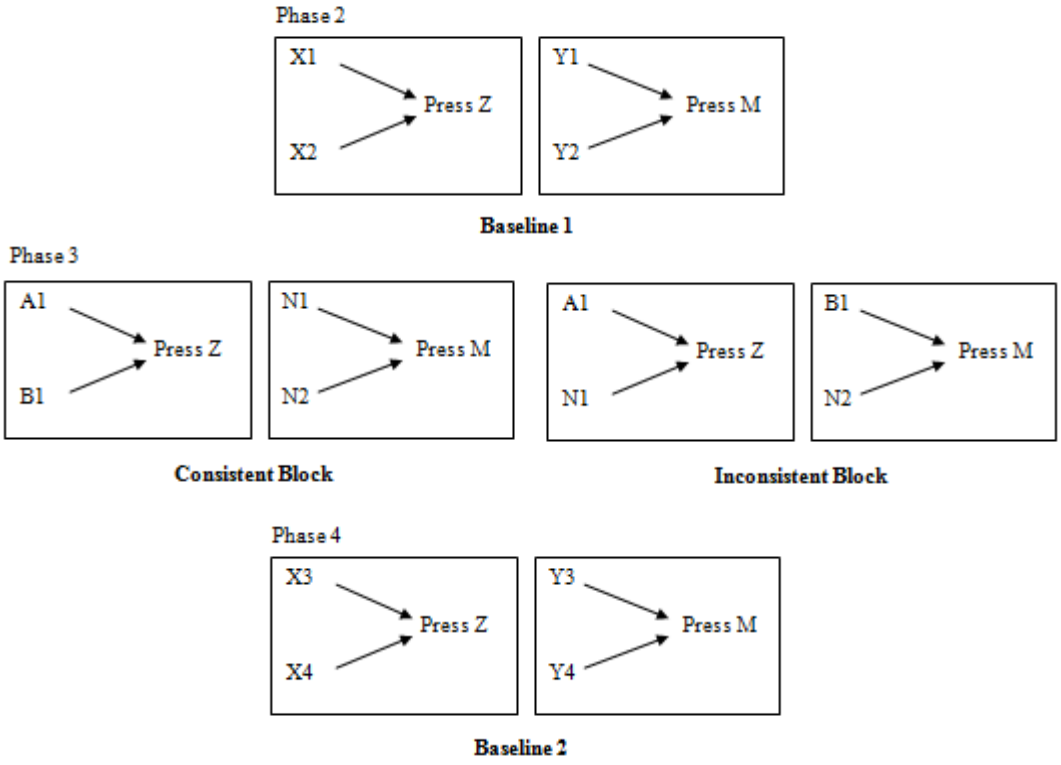

Figure 2.1.1: A sequential schematic representation of the response functions assigned to stimulus categories for Phases 2, 3, and 4 of the FAST. Note: adapted from O'Reilly et al. 2012.

Phase 2 and 4: Baseline blocks

Immediately after completing the practice block, participants were presented with the first of the two baseline blocks, during which relations between neutral, previously unrelated stimulus categories were again established through reinforced shared response functions. The aim of the two baseline blocks was to investigate the learning rate for the participant of acquiring a shared response function for two categories that were previously unrelated. In other words, the baseline blocks provided a baseline level of response acquisition against which the acquisition rates for common response functions across the two target categories can be compared. The presence of a second baseline block (block 4) merely served to

enhance the stability of the baseline score by providing a second measure at a separate moment in time. This afforded the opportunity to calculate average baseline scores and to examine practice effects for the test format itself (i.e., across the two baseline blocks within a single testing session). Prior to beginning the block participants were presented with the following instructions:

> *In the following section your task is to learn which button to press when a word appears on screen. IMPORTANT: During this phase you should press only the Z key or the M key. Please locate them on the keyboard now. This part of the experiment will continue until you have learned the task and can respond without error. To help you learn you will be provided with feedback telling you of you are right or wrong. If you have any questions please ask the researcher now. Press any key when you are ready to begin.*

During the first baseline block, the stimulus categories X1 (countries) and X2 (birds) shared the response key 'Z', while categories Y1 (sports) and Y2 (occupations) shared the response key 'M', while during the second baseline block, the stimulus categories X3 (trees) and X4 (dishes) shared the response key 'Z', while categories Y3 (fish) and Y4 (numbers) shared the response key 'M'. Both baseline blocks continued presenting stimuli in a randomised order until participants reached fluency, or until they reached 200 trials.

Phase 3: Testing blocks

Phase 3 of the FAST consisted of the two critical test blocks, which together index the strength of the A-B relation between the two target stimulus categories relative to two previously novel and unrelated 'N' categories. Prior to beginning both of the testing blocks, participants were presented with the following instructions:

> *In the following section your task is to learn which button to press when a word appears on screen. IMPORTANT: During this phase you should press only the*

Z key or the M key. Please locate them on the keyboard now. This part of the experiment will continue until you have learned the task and can respond without error. To help you learn you will be provided with feedback telling you if you are right or wrong. If you have any questions please ask the researcher now. Press any key when you are ready to begin.

Phase 3 consisted of both a consistent (i.e., where a shared response function is reinforced ACROSS the two categories of interest) and an inconsistent block (i.e., where a shared response function is reinforced BETWEEN each of the two the target categories and each of two additional categories). The blocks were presented in a random order within Phase 3. Specifically, in the inconsistent block, the stimulus categories A1 (descriptive words for Gay people) and N1 (colours) shared the reinforced response key press of the letter Z, while categories B1 (positive feeling words) and N2 (days of the week) shared the response key M. In the consistent block, the stimulus categories A1 (descriptive words for Gay people) and B1 (positive feeling words) shared the response key Z, while categories N1 (days of the week) and N2 (colours) shared the response key M. The pre-existing relational strength between the Gay words and Positive words was assessed by examining whether acquiring a common response function for the two target stimuli in the consistent block is facilitated or retarded relative to the rate of learning to criterion in the inconsistent block and to the two baseline blocks.

2.1.2.5.3 The Modern Homonegativity Scale

The MHS was administered to participants upon completion of the FAST procedure. With regard to why it was administered after rather than prior to the FAST, Nosek, Greenwald and Banaji (2005) found that whether the self-report scale is administered before or after the implicit measure exerts little influence over outcomes. However, these researchers do suggest that administering the self-report scale following the implicit measure may be

beneficial given that it does not alert participants to the specific nature of the test. Consequently, the MHS explicit measure was administered after the FAST, rather than before it.

2.1.2.5.4 Debriefing

Participants were fully debriefed regarding the nature of the experiment immediately upon completion, and were thanked for their participation. Additionally, at this point participants were invited to put forward any questions or queries regarding the research. Any questions were answered as comprehensively as possible, though follow-up also included the availability of a research supervisor in the Department of Psychology, at NUI Maynooth. Participants were also informed at this stage that all data analyses were to be conducted at a group level, and that no personally identifying information relating to them individually would be stored at any time.

2.1.3 Results

2.1.3.1 Modern Homonegativity Scale Scores

All participants in the current experiment were given an overall MHS score (Table 2.1.2), on the basis of which they were assigned either to Group One (low levels) or two (moderate-to-high levels), which was divided around the median MHS score of 28. Scores varied within the two groups, with more variance witnessed in Group One relative to Group Two. MHS scores in Group One ranged from 15 to 27 (M = 20.90, SD = 3.65), with scores tending to cluster around the high end and with some extremes. Scores in Group Two ranged from 28 to 41 (M = 31.45, SD = 3.38), with scores clustering around the low end and with fewer extreme cases.

Table 2.1.2: MHS scores for Group One and Group Two

Group One		**Group Two**	
Participant	MHS	**Participant**	MHS
2	17	**1**	28
4	27	**3**	30
5	15	**7**	28

6	23	**9**	31
8	23	**13**	34
10	22	**20**	28
11	16	**21**	32
12	21	**22**	31
14	24	**25**	30
15	23	**26**	28
16	16	**29**	36
17	16	**30**	30
18	19	**31**	30
19	20	**33**	36
23	25	**34**	41
24	24	**35**	34
27	18	**37**	28
28	20	**38**	32
32	26	**39**	30
36	23	**40**	32

2.1.3.2 Practice Blocks

As previously mentioned, the practice block (consisting of a fixed 16 trials) required no fluency criterion, and was intended to function only as a procedure to familiarise participants with the general test format. As such, this particular data will not be discussed here.

2.1.3.3 Baseline Blocks

Table 2.1.3 documents the number of trials required by each participant to reach the fluency criterion (10 correct responses in a row without error) for both baseline blocks. An inspection of the data set below indicates that within-participant acquisition rates typically decreased from the first to the second baseline block. Indeed, a Wilcoxin signed rank test revealed a statistically significant reduction in acquisition rates from the first to the second baseline block ($z = -2.185$, $p < .031$, two-tailed), with medium effect size, suggesting that participant are improving significantly in terms of performance as they progress through the procedure.

It was also important to factor into the current analysis potential age-related differences in mean baseline acquisition rates for participants, given the variance in the

current sample. In order to assess whether age influenced baseline acquisition rates, the sample was divided around the approximate median age (51 years of age) and a Mann-Whitney U test was conducted on mean baseline scores between those under 51 (n = 20, M = 28.4, SD = 9.4) and those over 52 (n = 20, M = 44.4, SD = 17). This analysis revealed a statistically significant difference in the number of trials to fluency in the expected direction between younger (Md = 27.3) and older (Md = 40.3) participants (z = -3.24, $p < 0.005$, r = .51). While it is unclear from the current data set, it may be possible that factors such as standard age-related cognitive decline or reduced familiarity with computerised tasks may account for this difference. Consequently, age will be duly factored into all subsequent analyses, and its potential impact upon FAST outcomes will be considered at a later stage.

Table 2.1.3: Number of trials to criterion across both baseline blocks (Phases 2 and 4). Participant numbers are listed in the column headed P.

Group One				**Group Two**			
P	Baseline 1	Baseline 2	Mean Baseline	**P**	Baseline 1	Baseline 2	Mean Baseline
2	41	26	33.5	**1**	33	48	40.5
4	31	14	22.5	**3**	27	30	28.5
5	26	28	27.0	**7**	20	13	16.5
6	13	25	19.0	**9**	68	37	52.5
8	43	20	31.5	**13**	14	47	30.5
10	34	16	25.0	**20**	19	30	24.5
11	16	12	14.0	**21**	31	44	37.5
12	20	18	19.0	**22**	86	137	86.0
14	42	38	40.0	**25**	33	30	31.5
15	36	19	27.5	**26**	60	26	43.0
16	15	30	22.5	**29**	30	24	27.0
17	61	19	40.0	**30**	25	38	31.5
18	55	30	24.5	**31**	74	33	53.5
19	26	31	28.5	**33**	33	95	64.0
23	68	38	53.0	**34**	69	57	63.0
24	32	35	33.5	**35**	94	57	75.5
27	23	32	27.5	**37**	57	21	39.0
28	66	17	41.5	**38**	52	46	49.0
32	113	26	26.0	**39**	54	21	37.5
36	45	40	42.5	**40**	32	20	26.0

2.1.3.4 Function Acquisition Speed Test

2.1.3.4.1 Acquisition Rates

It was expected that for Group Two (moderate-to-high homonegativity) it was expected that a larger number of trials would be required relative to Group One for the consistent block relative to the inconsistent block, given their supposed pre-existing history of relating Gay terms with Negative terms (i.e., not with positive terms). To test this prediction, it was necessary to investigate whether or not there was a significant difference at a group level between acquisition rates on the two critical test blocks for each of the two groups of participants. Thus, a Wilcoxin Signed Ranks test was conducted for both Group One and Group Two comparing the acquisition rates between the inconsistent and consistent blocks. The acquisition rate differential was found to be non-significant for both Group One (z = -.624, p > .05, two-tailed) and Group Two (z = -1.389, p > .05, two-tailed), indicating no FAST effect at the group level.

Furthermore, a closer inspection of the data set indicates that nearly half of the participants reported effects that were actually inconsistent with the current hypotheses. While some variance was expected for in Group One, it was nevertheless hypothesised that there would be a general trend towards predominantly higher trial requirements (signifying greater behavioural resistance) for the consistent block compared to the inconsistent block, for Group Two, with a general trend expected toward the reverse pattern in Group One. An inspection of the data set below, however (Table 2.1.4), revealed a somewhat perplexing pattern, with nearly half of participants within both groups showing effects in an unexpected direction (see Figure 2.1.2).

Table 2.1.4: Number of trials to criterion across both FAST blocks (Phase 3).
Note: the difference is obtained by subtracting the number of trials to criterion for the consistent block from the inconsistent block.

Group One				Group Two			
P	Inconsistent	Consistent	Difference	**P**	Inconsistent	Consistent	Difference
2	15	30	-15	**1**	20	25	-5
4	35	41	-6	**3**	93	24	69
5	27	12	15	**7**	47	16	31
6	30	44	-14	**9**	30	26	4
8	28	24	4	**13**	10	18	-8
10	30	24	6	**20**	38	32	6
11	25	25	0	**21**	33	48	-15
12	26	30	-4	**22**	38	51	-13
14	22	37	-15	**25**	19	43	-24
15	21	12	9	**26**	51	59	-8
16	13	77	-64	**29**	36	36	0
17	42	76	-34	**30**	17	39	-22
18	55	18	37	**31**	40	30	20
19	18	26	-8	**33**	52	22	30
23	39	19	20	**34**	62	37	25
24	24	46	-22	**35**	92	42	50
27	32	22	10	**37**	50	15	35
28	59	34	25	**38**	31	22	9
32	87	13	74	**39**	61	48	13
36	100	19	81	**40**	42	45	-3

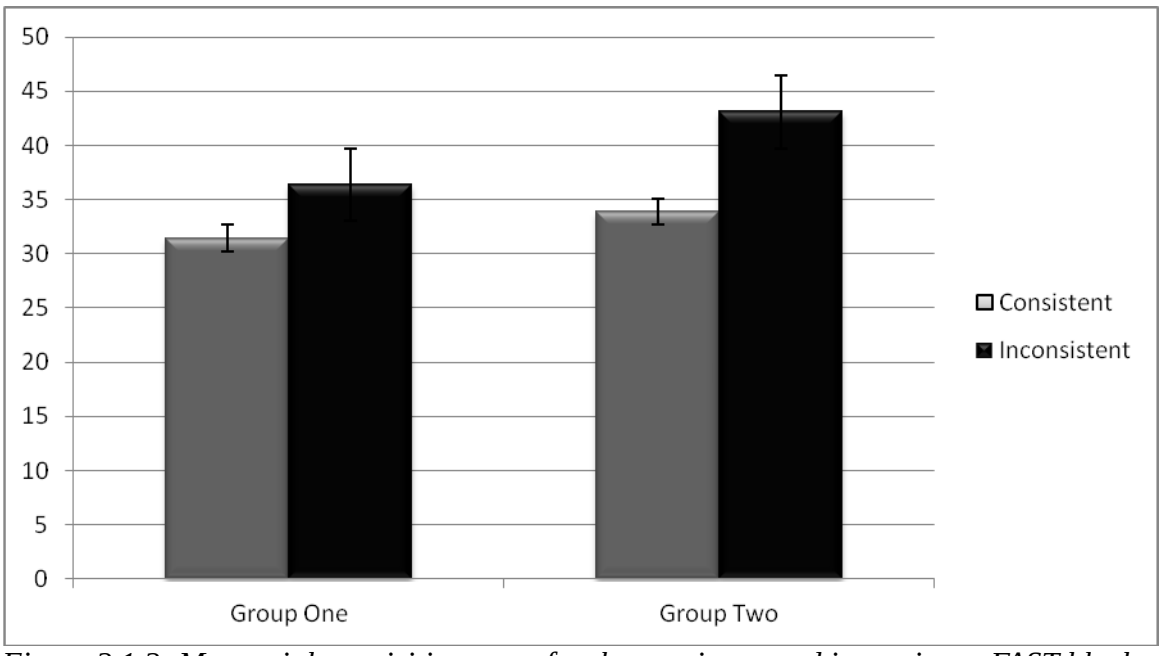

Figure 2.1.2: Mean trial acquisition rates for the consistent and inconsistent FAST blocks.

2.1.3.4.2 Strength of Relation Indices

To assess the strength of the pre-existing A1-B1 relation (in this instance between 'Gay' and 'Good'), a Strength of Relation (SoR) Index (O'Reilly et al., in press) was calculated for each participant. The SoR index is used to assess the strength of relation between the two target categories relative to two previously neutral stimulus categories when considered in relation to the natural logarithm of the mean baseline acquisition rates. It is calculated by dividing the difference in the acquisition rates recorded for the consistent and inconsistent blocks by the natural logarithm of the participant's mean baseline score. The log function serves merely to normalise data distribution by reducing skew and kurtosis. In this way, a large positive SoR index is indicative of higher strength stimulus relations, such that acquiring a common A1-B1 response function is facilitated relative to their baseline acquisition rates for previously unrelated stimulus categories. In contrast, a zero or negative SoR index reflects a lower strength or absent stimulus relations, such that acquiring a common A1-B1 response function is not facilitated or even retarded relative to their baseline acquisition rates. Thus, in the current experiment, large positive SoRs can be assumed to be indicative of a strong pre-existing Gay-Good (i.e., pro-gay) relation, while zero or negative SoRs can be assumed to be indicative of a weak or absent pre-existing Gay-Good relation.

Table 2.1.5 reveals a somewhat random distribution of SoR indices across the two groups, with no apparent trend emerging upon inspection. Single-sample t-tests revealed SoR indices were not significantly different to zero for the entire sample $t(39) = 1.478$, $p = .147$, as well as for Group One, $t(19) = .636$, $p = .446$, two-tailed, and Group Two $t(10) = 1.637$, $p = .174$, two-tailed, suggesting no overall FAST effect. Further to this, a Mann-Whitney U test comparing SoR indices between Group One ($n = 20$, $M = 3.19$, $SD = 22.42$) and Group Two ($n = 20$, $M = 5.78$, $SD = 15.8$) revealed no statistically significant difference in SoR values between the two groups ($z = -4.19$, $p > 0.05$, two-tailed), with Group Two (moderate-

to-high explicit homonegativity) reporting an unexpectedly higher average Gay-Good SoR index than Group One (i.e., greater behavioural resistance to the formation of a Gay-Good stimulus class). To gain a more comprehensive understanding of SoR distributions within the two groups, a simple Spearman's Rho correlational analysis was conducted to examine the relationship between the MHS and SoR indices for the sample as a whole, as well as for the individual groups. This analysis revealed a weak, non-significant correlation for the entire sample (r = .035, p = .832), and for SoR indices for Group One (r = -.007, p = .977) and Group Two (r = .108, p = .651). The scatter plot below (Figure 2.1.1) provides a clear graphical representation of the weak relationship between MHS scores and SoR indices within this experiment, with no discernible linear pattern within the graph.

Table 2.1.5: Strength of Relation (SoR) Indices. Note: large positive SoRs indicate a strong pre-existing Gay-Good *relation, while zero or negative SoRs suggest a weak or absent pre-existing* Gay-Good *relation*

Group One		**Group Two**	
Participant	[SoR]	**Participant**	[SoR]
2	-3.28	**1**	-3.11
4	-4.44	**3**	47.45
5	10.48	**7**	25.47
6	-11.02	**9**	2.32
8	-2.68	**13**	-5.39
10	4.29	**20**	4.32
11	.00	**21**	-9.55
12	-1.29	**22**	-6.73
14	-9.36	**25**	-16.10
15	6.25	**26**	-4.90
16	-47.40	**29**	.00
17	-21.25	**30**	-14.76
18	26.81	**31**	5.81
19	-5.52	**33**	16.60
23	11.62	**34**	13.96
24	-14.47	**35**	26.73
27	6.99	**37**	22.01
28	15.52	**38**	5.32
32	52.48	**39**	8.25
36	50.00	**40**	-2.12

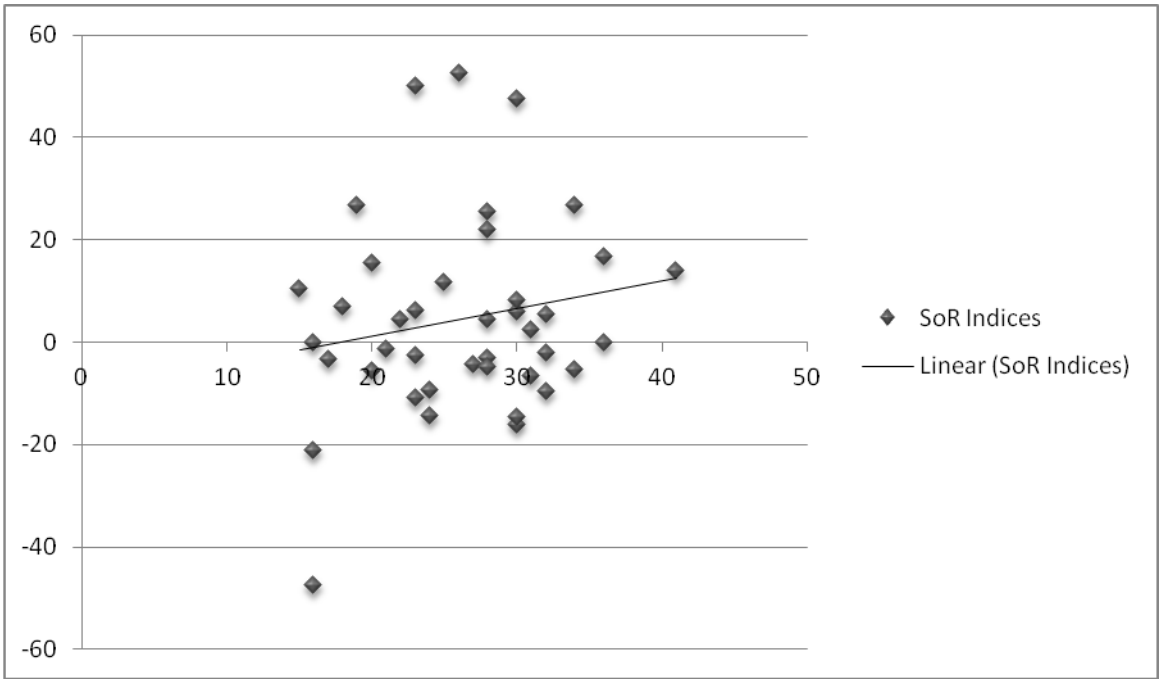

Figure 2.1.3: MHS-SoR Correlations for the Entire Sample

2.1.3.4.3 Order Effects

It was also important to assess whether or not the order of the two test blocks (i.e., whether the inconsistent or consistent block was presented first) impacted significantly upon outcomes. A Mann-Whitney U test conducted for the entire sample indicated that trial requirement differentials across FAST blocks were not significantly different based on the order in which they were administered (z = -7.83, p = .434, two-tailed), suggesting that order effects are not observed in the current data set.

Table 2.1.6: Running order for inconsistent and consistent test blocks (1: Consistent first, 2: Inconsistent first, P refers to Participant)

Group One																				
P	2	4	5	6	8	10	11	12	14	15	16	17	18	19	23	24	27	28	32	36
Order	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2
Group Two																				
P	1	3	7	9	13	20	21	22	25	26	29	30	31	33	34	35	37	38	39	40
Order	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2

2.1.3.6 Summary

This experiment revealed a significant reduction in baseline acquisition rates across the two baseline blocks for all participants (i.e., it revealed a practice effect). Additionally, age contributed significantly to the number of trials required to reach criterion in the baseline blocks, with increases in age generally associated with increases in acquisition rates. Secondly, with regard to the two FAST testing blocks, the acquisition rate differentials between the inconsistent and consistent blocks were non-significant for both groups, suggesting no overall FAST effect. Further to this, over half of all participants in the sample showed acquisition rate differentials in a direction inconsistent with MHS scores. Similarly, SoR distribution was somewhat random, with no significant differences found between SoRs for Group One and Group Two. Additionally, Group Two were found to have unexpectedly higher average homopositivity SoRs than Group One. Correlational analyses reported a weak association between MHS and SoR indices for the sample as a whole as well as for the two groups. Finally, this experiment also revealed that the order in which the two testing blocks were administered did not affect acquisition rates for the two FAST blocks.

2.1.4 Discussion

The aim of the current experiment was to assess the suitability of multi-exemplar verbal stimuli in a Homopositivity FAST. The effectiveness of the verbal stimuli was to be assessed primarily through the use of a "known-groups" approach, wherein two distinct groups were employed based on participants' self-reported attitudes toward Gay individuals. It was conjectured from the available literature that, while self-presentational concerns may distort self-reports from those claiming strong pro-gay attitudes (ascertained from the MHS scale), it might be possible to identify those who are explicitly anti-gay using this approach. Consequently, it was hypothesised that, based on their verbal histories, FAST scores for Group Two (moderate-to-high explicit homonegativity) would be notably different to those in Group One (low explicit homonegativity). It was also conjectured from previous research that there would be a general trend in the data toward weak or reversed Gay-Good relations (i.e., that there would be an overall FAST effect), given the overall anti-gay bias that would be expected in a typical Irish population sample.

Statistical analyses conducted above suggested no logical or discernible trends in the FAST scores of the current sample. Firstly, the analyses indicated no overall FAST effect in any direction (i.e., no particular bias towards being pro- or anti-gay), either for the sample as a whole or for the individual groups. In addition, there was no significant correlation between MHS scores and SoR indices. Compounding this further was the finding that over half of all participants performed in way that directly opposed the hypothesised outcomes based on self-reported levels of homonegativity. In short, it would appear the current experiment yielded none of the predicted effects.

The somewhat random nature of the data set makes it difficult to interpret the current findings. This experiment utilised four widely known descriptive words for Gay individuals, which were chosen based on their perceived semantic representativeness of the categories

'Gay' and 'Good'. Verbal exemplars were chosen because, as previously discussed, research has suggested a distinction between semantic-verbal stimuli, which are believed to have high levels of representation, and perceptual-pictorial stimuli, which are considered to have comparatively lower levels of representation (Marsolek, 1999; Tanaka et al., 1999; Whatmough et al., 2004). According to this distinction, verbal stimuli may have been more appropriate in representing the target category 'Gay', given that these exemplars should have instantiated the wider concept more easily than more specific pictorial exemplars. However, judging on the above outcomes, one possible interpretation of the above results is that perhaps semantic-verbal stimuli may not be suitable for attitudes toward Gay individuals.

There are several reasons as to why this may have been the case. Firstly, it may be true that while the verbal stimuli are semantically representing the target category, they are failing to evoke the target category in the desired way (i.e., a Gay individual). The importance of selecting exemplars that accurately instantiate the super-ordinate stimulus category has been widely noted in the literature (e.g., Lane et al., 2007), with numerous studies revealing the influence context can have in how participants evaluate the overall construct (e.g., 'Gay'; De Houwer, 2001; Mitchell et al., 2003; Nosek et al., 2005). For example, IAT research has examined how prejudicial attitudes to Black people can vary depending on the context within which they are presented (e.g., well liked Black athletes versus disliked White politicians; Mitchell et al., 2003). This research may inform the current findings here. Firstly, though the words that were chosen as exemplars for the category 'Gay' were selected from the LGB literature (e.g., *Gay, Homosexual,* etc.), given they are also frequently used among the general population, it may be possible that this modulated how participants evaluated the overall construct. It has been suggested that exemplars should be selected that are unambiguous members of the super ordinate category, in order to eliminate any potential cross-class overlap (De Houwer, 2001; Steffens & Plewe, 2001; Lane et al.,

2007). That is, exemplars should be selected that are clear and unambiguous, and that do not contain conflicting stimulus functions (i.e., the features of the stimuli that control behaviour) which override the target construct. However, one possible interpretation of the above findings could be that the familiar nature of the words among the general public in Ireland altered the super ordinate class of exemplars 'Gay' from depicting 'Gay individuals', to the comparatively neutral stimulus class 'words I hear frequently'. Thus, in this way, the familiarity of the words would have transformed how participants engaged with and evaluated the stimulus category.

Another possibility could be that, while the exemplars 'Queer', 'Homosexual' and 'Gay' were selected from the LGB literature as non-intentionally offensive words, they may potentially still hold strong pejorative connotations for the general public. This may be especially true for the word 'Gay', which is a particularly widely used pejorative term in Irish society (Norman, 2006; Mac an Ghaill et al., 2006). Therefore, in this instance, rather than purely exemplifying same-sex relationships, the perceived negative functions of the words would have been the defining feature of the stimulus class. Thus, it is possible that while the words may have been semantically descriptive of the target category, they did not instantiate the construct in the required way. As previously discussed, it has been suggested that verbal exemplars may instantiate a construct more widely than images by generating a wider field of information related to that particular construct (Foroni & Bel-Bahar, 2010). However, this raises some interesting questions regarding the use of verbal exemplars in implicit testing research, given that, as can be seen even in the case above, words may contain a variety of conflicting stimulus functions depending on the context. Consequently, it may be more difficult to retain a high degree of control over verbal exemplars than images, though this would need to be explored further.

An additional way in which the frequent exposure to such terms may have confounded results could be that it possibly eliminated the affective element of the exemplars. It has been long argued that attitudes, and particularly prejudicial attitudes, have a strong affective component (e.g., Fleming, 1967; Ostrom, 1969; Millar & Tesser, 1990). Thus, some attitudes, such as anti-gay attitudes, may be best conceived as emotional responses to Gay people, rather than reasoned verbal responses (Farley, 2000). In other words, while verbal relations are surely involved in how we form and relate attitudes to others, they often involve powerful emotional components that form the basis of a strong evaluative response to the relevant exemplars (see Haddock & Zanna, 1993; Stangor, Sullivan & Ford, 1991, for reviews). Put simply, while an individual may have worked out various reasons to dislike Gay people, which they may state verbally, their dislike is unlikely to be entirely rational or "intellectual". From a behaviour-analytic point of view, this simply means that there are likely emotional as well as other response functions (e.g., labels) transforming in accordance with the verbal network that constitutes the attitude. That is, while there may be little resistance to the formation of verbal classes that are incongruous with the attitude, there may be much more behavioural resistance to the formation of functional stimulus classes involving emotive stimuli that are incongruous with the functional classes of emotional stimulus functions that form part of the attitude. Indeed, previous behaviour analytic work has examined the interaction between functional classes of emotional stimuli and purely verbal classes of stimuli and has suggested that attitudes are composed of networks containing stimulus relations of both types (see Roche, Barnes & Smeets, 1997). Thus, as is now becoming apparent from the current findings, and as published research suggests (e.g., Hudepohl et al., 2010; Kimmel, 1997; Mahaffey, Bryan & Hutchison, 2005; Parrott, 2008), anti-gay attitudes may be quite evaluative rather than cognitive in nature, and thus semantic-verbal representation, while clear and unambiguous,

may not be suitable for use in implicit tests for this particular kind of prejudice (i.e., in order to produce strong behavioural resistance to stimulus class formation).

In addition to the issues with the verbal stimuli, another important finding in this experiment was the influence of age on baseline acquisition rates. Older participants required significantly more trials than younger participants to reach fluency criterion in the baseline blocks, suggesting that age may exert a negative influence on FAST performance. However, this is not unexpected given the widely noted influence of age on computer-based tasks, which is commonly attributed to reduced experience with computers in older age groups (e.g., Czaja & Sharit, 1993; Ownby, Czaja, Loewenstein & Rubert, 2008). Consequently, given that age differences are not the primary focus of the current research, it may be necessary in future experiments to focus on a more homogenous sample in terms of age to minimise any confounding influences of extraneous variables on FAST scores.

An additional important finding in the current experiment relates to the practice effect observed in the baseline phases. There was a significant reduction in the number of trials required to achieve fluency from the first to the second baseline blocks, suggesting that participants' performance is significantly improving as the progress through the procedure (i.e., the FAST is subject to 'practice effects'). Research has suggested this issue also affects the IAT, with some studies revealing that participants can become notably faster and more accurate with repeated exposures to the same IAT (Fiedler & Bluemke, 2005; Röhner, Schröder-Abé, Schütz, 2011). Indeed, a relatively recent meta-analysis revealed the average test-retest reliability coefficient of the IAT (i.e., the stability of participants' scores on the IAT after repeated administrations) to be .56 (Nosek, Greenwald & Banaji, 2007), which may be considered somewhat low given that accepted values are usually around .80 or higher (e.g., Aiken, 1994). Therefore, based on this research, the IAT does appear to be subject to pronounced practice effects, given that almost half of the IAT effect vanishes after one

administration. However, while this is undoubtedly a significant finding and will be borne in mind throughout this research programme, it is important not to create confounds by altering too many variables at once in subsequent experiments. That is, it is of considerable importance to ensure that the core issue of stimulus selection has been sufficiently addressed in order to produce a prototype that can thence honed in relation to stability of effects. Therefore, the issue of practice effects will remain an observation until the issue of stimuli has been addressed successfully.

In conclusion, the current experiment documented none of the predicted effects, with no identifiable trend or pattern exhibited in the data set. Had the MHS scores recorded for participants at least correlated at face value with SoR scores, we might conclude that the surprising outcome was due to was a lack of attitudinal bias in any direction and an overlay narrow range of scores on the MHS. However, this was not the case. Consequently, it may be logical to infer here that the stimuli were unsuccessful in accurately representing the target categories. This may have been the result of the verbal stimuli not being emotionally evocative for this particular attitude, or from cross-category overlap. Thus, this experiment has highlighted the importance of depicting the super ordinate stimulus categories in the desired way. Research suggests that anti-gay attitudes are quite evaluative rather than cognitive in nature and, therefore, it may be possible that semantic-verbal representation is not appropriate for a FAST examining anti-gay prejudice. Consequently, Experiment 2 will focus on the relative effectiveness of pictorial stimuli in a Homopositivity FAST.

2.2 Experiment 2

2.2.1 Introduction

The current experiment aims to build on the findings of Experiment 1 and investigate the suitability of multi-exemplar pictorial stimuli for a Homopositivity FAST. As discussed previously, an important distinction exists in the IAT literature between pictorial and verbal stimuli. Typically, picture-IATs are associated with less extreme effect sizes than word-IATs (see Foroni & Bel-Bahar, 2010) and, consequently, the first experiment in this thesis investigated the relative suitability of verbal stimuli in a Homopositivity FAST. This was assessed primarily through the use of a 'known-groups' approach, wherein FAST scores from participants with high self-reported homonegativity were contrasted with the scores from the rest of the sample. As suggested previously, the words utilised in Experiment 1 may have been unsuitable exemplars for the target categories, given that they failed to engender any differences in FAST performances across test blocks for any sample strata (e.g., those with high-self reported homonegativity).

The current experiment will now focus on the utility of pictorial stimuli and their effectiveness in a Homopositivity FAST. Utilising pictorial stimuli may be an improvement upon the first experiment for numerous reasons. Firstly, pictorial exemplars are argued to be more perceptual than verbal exemplars (Henson et al., 2002; Ito & Ulrand, 2005; Rossion et al., 2003; Todorov et al., 2007), which, as outlined in detail in the discussion section above, may make them more effective than words at depicting a construct with a strong affective component (e.g., Huijding and de Jong, 2005). Moreover, compared to verbal stimuli, some researchers has argued that pictorial stimuli contain more concept-relevant information, which may make them more appropriate for representing prejudicial constructs (Gschwendner, Hofmann, & Schmitt, 2008).

Another reason pictorial stimuli may prove more suitable for a Homopositivity FAST is that they may be more ecologically valid exemplars of the construct of interest (i.e., Gay individuals) than verbal stimuli. It has been argued that not all relevant concepts can be adequately instantiated by a specific word, even though they may be more semantically representative exemplars of the stimulus category (Huijding & de Jong, 2005). In other words, although descriptive words depicting the target construct may be appropriate semantic exemplars, the attitudes held by prejudiced individuals toward the *words* 'Gay' or 'Homosexual' as opposed to an actual Gay individual could be entirely different. In the context of phobia research, for example, the word 'spider' is considered to be a excellent semantic description of the particular construct of interest for spider phobia (i.e., of the class of spiders). However, individuals with spider phobia are phobic of *actual* spiders rather than the word itself, and unless fear has generalised to include even descriptors of the phobic stimulus class, many spider-phobic individuals would not respond to the word spider in the same as they will to an actual spider, or a picture of a spider (e.g., Teachman et al., 2001. In the same way, for a concept such as prejudicial anti-gay attitudes, it may be necessary to utilise more visually perceptual stimuli that evoke the emotional stimulus functions of interest (e.g., hate, disgust) associated with the stimulus class.

The current experiment will investigate the use of pictorial stimuli depicting Gay individuals as exemplars of the category 'Gay'. It is important, however, that the individuals are evaluated instantly as belonging to that stimulus class. For this purpose, images of widely known public figures for whom their sexual orientation is a defining character feature will be utilised. This technique has been used previously in the implicit testing literature, whereby public figures or celebrities have been employed as exemplars of a particular stimulus category (e.g., politicians, athletes, musicians, etc.) on the basis of their obvious membership to a particular social group (e.g., Brunel, Tietje & Greenwald, 2004; Mitchell et al., 2003).

However, it is of considerable importance to also clearly establish the primary focus of the image exemplars to be employed (i.e., whether they are images of men or women, or a combination of both). As previously discussed, research has identified an important distinction between prejudicial attitudes towards Gay men and Lesbian women (e.g., Herek, 1984), revealing that greater modern homonegativity and anti-gay aggression tends to be directed toward Gay men rather than lesbian women (Breen & Karpinski, 2013; Morrison & Morrison, 2011). Indeed, one study comparing implicit attitudes toward Gay and Straight people revealed how the inclusion of images of Lesbian women as stimuli for the category 'Gay people' resulted in weaker implicit pro-straight bias than when similar images of Gay men were employed (Nosek et al., 2005). While this is not to suggest that prejudice and discrimination against lesbian women is any less pronounced at a societal level, it would appear there is tendency toward more visceral or emotional responses to images of Gay men relative to Lesbian women. Therefore, in order to narrow the focus and minimise the potential confounding effects of various stimulus types, the current experiment will exclusively utilise images of widely known Gay men.

Experiment 2 will be identical in format to Experiment 1, with the only differences being that pictorial rather than verbal stimuli will be used, and that a more homogenous participant sample in terms of age will be employed. As such, the aims and expectations will be the same. The effectiveness of these pictorial stimuli will again be assessed largely through the use of a group-based 'known-groups' paradigm, wherein FAST scores from participants with high levels of explicit homonegativity (ascertained by the MHS) will be contrasted with the rest of the sample. It is expected that there may be general patterns visible within the data based on MHS scores. More specifically, it is expected that strongly anti-gay individuals will encounter behavioural resistance to forming a functional response class consisting of stimuli with 'Gay' and 'Good' response functions. In contrast, a more

varied outcome in terms of trial requirements across FAST blocks is expected for those with lower explicit levels of anti-gay attitudes. Finally, the current experiment again expects to observe a general trend toward anti-gay bias at the group level, as detected by the current version of the Homopositivity FAST.

2.2.1 Method

2.2.2.1 Participants

Participants for this experiment were recruited from either the NUI, Maynooth undergraduate population or from acquaintances of the experimenter via word of mouth recruitment. The 36 participants were all adults (over 18 years of age) who did not suffer from any condition that would interfere with their ability to perform routine cognitive tasks. The current experiment imposed an exclusion criterion identical to Experiment 1 for those participants who failed to reach fluency within 200 trials on the first baseline block. Consequently, six participants were excluded from the data set. Of the 30 participants remaining (14 male and 16 female), all were aged between 18 and 23 (M = 20.33, SD = 1.241)[2].

2.2.2.2 Ethical Considerations

As discussed previously, all elements of this research process received prior approval from the NUI, Maynooth Ethical Committee, and were all in accordance with the Psychological Society of Ireland (PSI) code of ethics, and with the NUI, Maynooth Department of Psychology guidelines for safe working practice in psychological research.

2.2.2.3 Materials

2.2.2.3.1 The FAST Software Package and Stimulus Set

The version of the FAST utilised in the current experiment was again developed with the computer programming software Livecode™, and presented on a 15” laptop (1920×1080

[2] Margaret Anne Brennan, an undergraduate student at NUI, Maynooth, assisted with data collection as part of her final year research project.

pixel resolution). Except for a small number of images specified below, pictorial stimuli for this experiment were obtained from the International Affective Picture System (IAPS) database (Lang, Bradley, & Cuthbert, 1999). Stimuli were matched roughly for standardised IAPS arousal ratings and valence across all blocks except for the consistent and inconsistent FAST blocks, in which one of the stimulus categories consisted of positively valenced and highly arousing images (as reported in the IAPS manual). Images were cropped slightly to achieve a perfectly square shape and in a small number of cases to eliminate peripheral objects in the image view. All images were 150x150 pixels in size (at a resolution of 72 dpi). During all test blocks, stimulus categories were comprised of four image exemplars. The alphanumeric category code (identical to Experiment 1), category description, and their corresponding IAPS image identifier numbers are listed below (Table 2.2.1).

Table 2.2.1: Pictorial exemplars employed in Experiment 2

Alphanumeric	**Stimulus Category**	**IAPS Image Number/Image Description**
Phase 1: Practice block		
P1	Babies	2040, 2050, 2250, 2260
P2	Nature	5201, 5220, 5760, 5780
Q1	Sky	5593, 5594, 5990, 5991
Q2	Zoo animals	1601, 1661, 1810, 1812
Phase 2: Baseline 1		
X1	Birds	1333, 1419, 1450, 1740
X2	Cakes	7200, 7250, 7270, 7282
Y1	Mushrooms	5531, 5532, 5533, 5544
Y2	Vehicles	7130, 7140, 7595, 8510
Phase 3: Test Block		
A1	Gay men	Boy George, Elton John, Graham Norton, George Michael

B1	Skiers	8021, 8030, 8034, 8190
N1	Chinese symbols	Four easily discriminable Chinese symbols
N2	Colours	Green, Purple, Red, Yellow
Phase 4: Baseline 2		
X3	Food	7281, 7291, 7350, 7480
X4	Furniture	7025, 7235, 7705, 7710
Y3	Men	2102, 2214, 2215, 2493
Y4	Household pets	1510, 1540, 1610, 1750

2.2.2.3.2 The Modern Homonegativity Scale

In order to categorise participants into either Group One (low explicit homonegativity) or Group Two (moderate-to-high explicit homonegativity), this experiment again employed the Morrison and Morrison (2002) Modern Homonegativity Scale (Appendix 2) as a measure of self-report levels of explicit homonegativity. The scale used is identical to the version used in the first experiment outlined above (i.e., the 10-item MHS measure of homonegative attitudes toward Gay men).

2.2.2.4 Design

The current experiment again employed a between-groups quasi-experimental design with homonegativity scores forming one independent variable and the two critical consistent and inconsistent block types of the FAST together forming a second variable controlling test performance.

2.2.2.5 General Experimental Sequence

2.2.2.5.1 Informed Consent

The procedure for informed consent was identical to Experiment 1.

2.2.2.5.2 Function Acquisition Speed Test

The FAST procedure adhered to the same general experimental format outlined in Experiment 1, with an identical four-phase sequence (i.e., a practice phase, two baseline phases, and a testing phase). The specific reinforcement contingencies for each phase of the FAST are outlined below.

Phase 1: Practice block

During this phase, relations were established between categories P1 (babies) and P2 (nature), which shared the response key M, and between categories Q1 (sky) and Q2 (zoo animals), which shared the response key Z.

Phase 2 and 4: Baseline blocks

During the first baseline block, the stimulus categories X1 (birds) and X2 (cakes) shared the response key 'Z', while categories Y1 (mushrooms) and Y2 (vehicles) shared the response key 'M', while during the second baseline block, the stimulus categories X3 (food) and X4 (furniture) shared the response key 'Z', while categories Y3 (men) and Y4 (household pets) shared the response key 'M'.

Phase 3: Testing blocks

During the inconsistent block, the stimulus categories A1 (images of well-known Gay men) and N1 (Chinese symbols roughly translating as 'nonsense') shared the response key Z, while categories B1 (positive images of people skiing) and N2 (colours) shared the response key M. In the consistent block, the stimulus categories A1 (images of well-known Gay men) and B1 (positive images of people skiing) shared the response key Z, while categories N1 (Chinese symbols) and N2 (colours) shared the response key M.

2.2.2.5.3 The Modern Homonegativity Scale.

As in Experiment 1, the MHS was administered to participants following the FAST procedure.

2.2.2.5.4 Debriefing

The same debriefing as employed in Experiment 1 was again employed here.

2.2.3 Results

2.2.3.1 Modern Homonegativity Scale Scores

All participants in the current experiment were given an overall MHS score (*M* = 27.30; *SD* = 8.24; Table 2.2.1), on the basis of which they were assigned either to Group One (low levels) or Two (moderate-to-high levels), which was divided around the median MHS score of 28. Scores varied within the two groups, with more variance witnessed in Group One relative to Group Two. MHS scores in Group One ranged from 15 to 27 (*M* = 20.13; *SD* = 3.204), with scores tending to cluster around the lower end and with some extremes. Scores in Group Two ranged from 28 to 42 (*M* = 34.47; *SD* = 4.673), with scores clustering around the higher end and with fewer extreme cases.

Table 2.2.2: MHS scores for Group One and Group Two

Group One		**Group Two**	
Participant	MHS	**Participant**	MHS
1	21	**2**	36
5	22	**3**	31
7	20	**4**	29
8	14	**6**	35
9	21	**17**	30
10	21	**19**	30
11	21	**20**	34
12	21	**22**	36
13	21	**23**	34
14	26	**24**	40
15	20	**25**	42
16	14	**26**	31
18	21	**27**	39
21	23	**28**	28
30	16	**29**	42

2.2.3.2 Practice Blocks

As addressed previously, the practice block served only to familiarise participants with the procedure and, as such, the practice block data will not be discussed here.

2.2.3.3 Baseline Blocks

Table 2.2.3 documents the number of trials required by each participant to reach the fluency criterion across the two baseline blocks. An inspection of the data set reveals a similar pattern to that observed in Experiment 1, with acquisition rates typically decreasing from the first to the second baseline block. A Wilcoxin Signed Ranks test revealed a statistically significant reduction in trial acquisition rates from baseline 1 to baseline 2 for the entire sample ($z = -3.106$, $p < 0.05$, two-tailed), with large effect size. Again, this indicates that participants' enhanced familiarity with the format of the FAST is impacting upon acquisition rates, making subsequent blocks easier as they progress through the procedure.

Table 2.2.3: Number of trials to criterion across both baseline blocks (Phases 2 and 4. Participant numbers are listed in the column headed P.

Group One				**Group Two**			
P	Baseline 1	Baseline 2	Mean Baseline	**P**	Baseline 1	Baseline 2	Mean Baseline
1	55	21	38.0	**2**	81	17	49.0
5	38	26	32.0	**3**	71	26	48.5
7	38	22	30.0	**4**	61	47	54.0
8	22	13	17.5	**6**	56	13	34.5
9	31	10	20.5	**17**	77	31	54.0
10	15	43	29.0	**19**	34	11	22.5
11	33	26	29.5	**20**	42	65	53.5
12	13	16	14.5	**22**	31	26	28.5
13	51	14	32.5	**23**	22	21	21.5
14	21	36	28.5	**24**	94	64	79.0
15	25	14	19.5	**25**	66	19	42.5
16	38	17	27.5	**26**	32	14	23.0
18	15	20	17.5	**27**	21	32	26.5
21	18	11	14.5	**28**	20	24	22.0
30	21	22	21.5	**29**	28	13	20.5

2.2.3.4 The Function Acquisition Speed Test

2.2.3.4.1 Acquisition Rates

With the group structure similar here as in the previous experiment (i.e., Group One consisted of those with low explicit homonegativity and Group Two of those with higher explicit homonegativity), it was again expected that for Group Two a larger number of trials would be required to reach fluency for the consistent relative to the inconsistent block relative

to Group One. A Wilcoxin Signed Ranks test assessing whether there was a significant difference at a group level between acquisition rates across blocks, revealed a non-significant differential between the consistent and inconsistent blocks for Group One (z = -5.40, p = .589), suggesting no FAST effect. However, for Group Two, a Wilcoxin Signed Ranks test revealed a statistically significant difference between acquisition rates for the two critical blocks (z = -2.201, p = .028). Interestingly, as evidenced by Figure 2.2.2, this difference was in the unexpected direction, with participants in this group typically requiring fewer trials to achieve fluency for the consistent block relative to the inconsistent block (indicating behavioural facilitation relative to the inconsistent block).

Table 2.2.4: Number of trials to criterion across both FAST blocks (Phase 3).

	Group One				**Group Two**		
P	Inconsistent	Consistent	Difference	**P**	Inconsistent	Consistent	Difference
1	18	13	5	**2**	49	52	-3
5	21	14	7	**3**	16	26	-10
7	22	27	-5	**4**	59	31	28
8	55	10	45	**6**	20	20	0
9	30	12	18	**17**	36	36	0
10	17	47	-30	**19**	21	10	11
11	29	45	-16	**20**	34	64	-30
12	12	73	-61	**22**	42	10	32
13	28	45	-17	**23**	38	16	22
14	27	13	14	**24**	42	15	27
15	24	13	11	**25**	26	13	13
16	13	17	-4	**26**	21	15	6
18	13	28	-15	**27**	71	17	54
21	12	10	2	**28**	61	10	51
30	30	39	-9	**29**	22	10	12

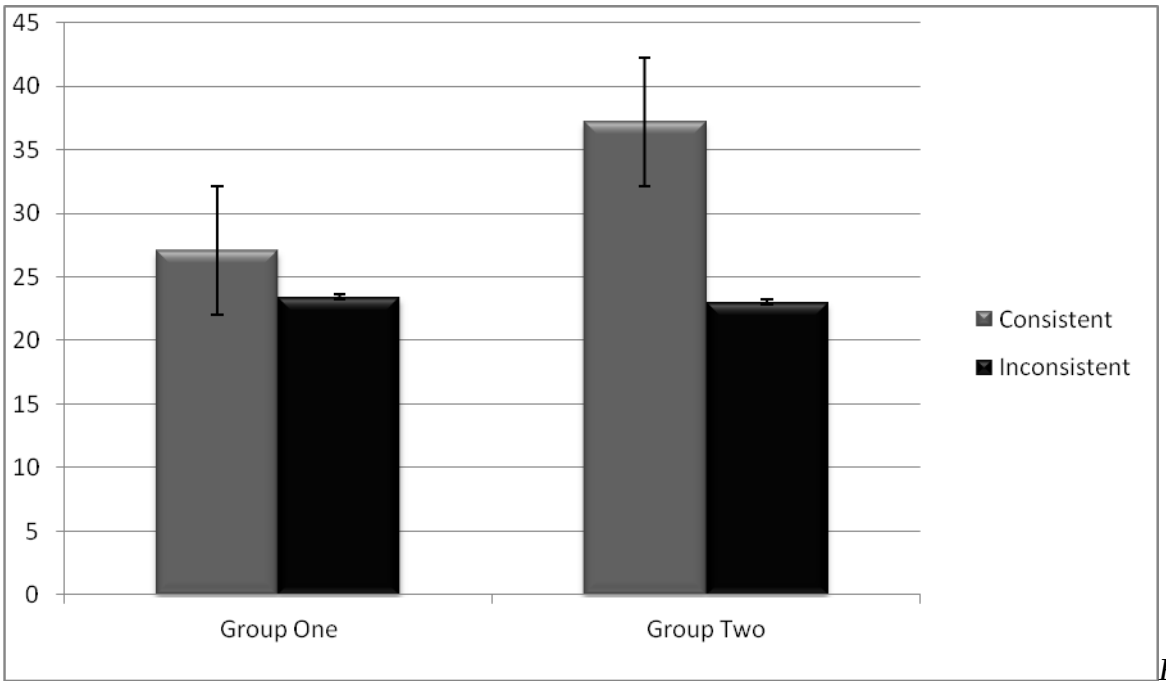

Figure 2.2.2: Mean trial acquisition rates for the consistent and inconsistent FAST blocks for Experiment 2

2.2.3.4.2 Strength of Relation Indices

This experiment again calculated a simple SoR index for each participant, which represented the strength of the pre-existing A1-B1 stimulus relation. It should be reiterated here that large positive SoRs indicate a strong pre-existing Gay-Good (i.e., pro-gay) relation, while zero or negative SoRs suggest a weak or absent pre-existing Gay-Good relation. Table 2.2.5 reveals a somewhat perplexing data set, with no logical pattern emerging upon inspection. Overall, while there is a tendency in the sample toward generally positive SoR indices (n = 30, M = 3.677, SD = 18.29), scores do cluster around the lower end, with a high number of extreme cases. Within Group One (n = 16, M = -2.84, SD = 19.33), SoRs tend to be clustered around the lower end with some extreme cases, while SoRs for Group Two (n = 16, M = 10.19, SD = 15.11) are unexpectedly clustered around the higher end, although again there are some extreme cases. Single-sample t-tests were conducted in order to investigate whether there was a significant overall bias within either group (i.e., whether SoR indices were significantly different to zero) and for the sample overall. These analyses revealed a

non-significant overall effect for the sample as a whole, t (29) = 1.101, p = .280, two-tailed, and within Group One, t (14) = -.568, p = .579, two-tailed, however, there was an unexpected significant positive trend for Group Two, t (14) = 2.163, p = .020, two-tailed. To investigate whether SoR indices differed significantly between the two groups, a Mann-Whitney U test was conducted, which revealed a statistically significant difference between SoR indices between the two groups (z = -2.053, p = .04, two-tailed) with medium effect size (r = .34). However, this difference was again in the unexpected direction, with Group Two demonstrating evidence of stronger A1-B1 relations than Group One. In addition, Spearman's Rho correlational analyses revealed a non-significant correlation between SoR indices and MHS scores for the sample overall (r = .337, p = .069), though this correlation was in a positive direction. Additionally, correlation analyses revealed a non-significant correlation between MHS scores and SoR indices within Group One (r = -.26, p = .926) and Group Two (r = -.998, p = .977), suggesting a relationship between FAST and MHS scores in the unexpected direction (i.e., increases in MHS scores correlating with higher Homopositivity SoR indices). The scatter plot below (Figure 2.2.1) also reveals this unexpected linear pattern to the data, which, although non-significant, does attest to the surprising positive correlation between the MHS scores and SoR indices for the sample overall.

Table 2.2.5: Strength of Relation (SoR) Indices for each participant

Group One		**Group Two**	
Participant	[SoR]	**Participant**	[SoR]
1	3.16	**2**	-1.78
5	4.64	**3**	-5.91
7	-3.38	**4**	16.18
8	36.30	**6**	.00
9	13.74	**17**	.00
10	-20.55	**19**	8.15
11	-10.89	**20**	-17.34
12	-52.59	**22**	16.54
13	-11.29	**23**	21.92
14	9.70	**24**	13.64

15	8.53	**25**	7.98
16	-2.78	**26**	4.41
18	-12.10	**27**	38.03
21	1.72	**28**	38.06
30	-6.77	**29**	13.00

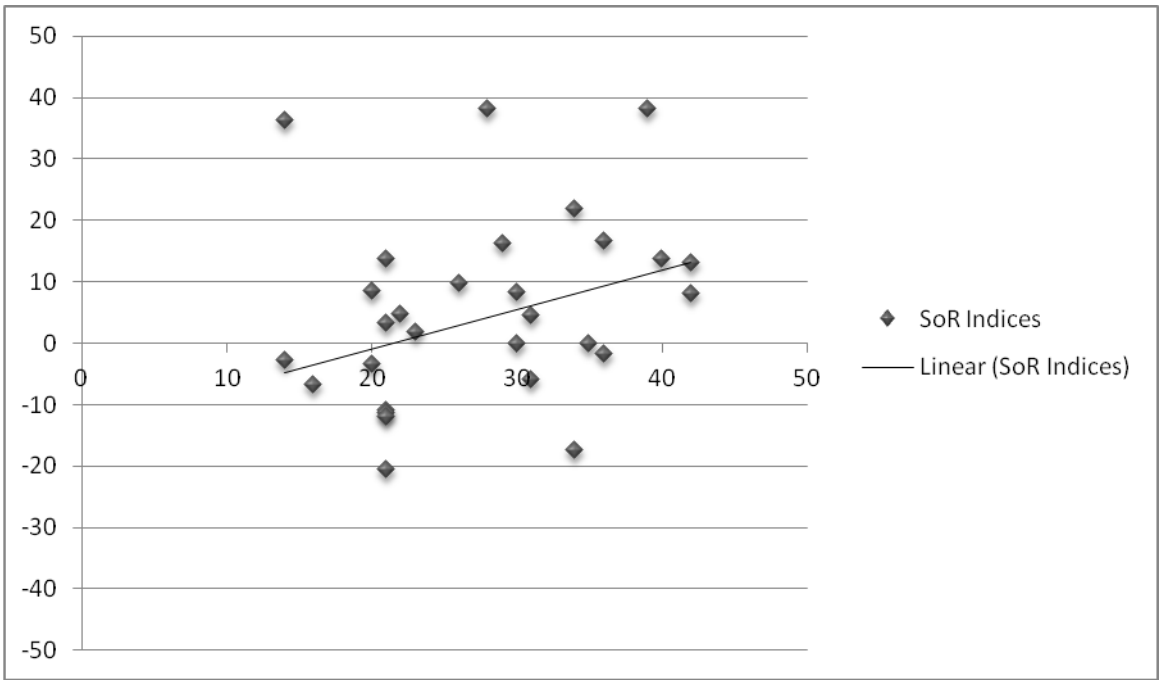

Figure 2.2.1: MHS-SoR Correlations for the Entire Sample

2.2.3.4.3 Order Effects

To investigate whether the order in which the FAST portion of the experiment was administered (i.e., whether the consistent or inconsistent block was presented first) influenced outcomes, a Mann-Whitney U test was conducted. This analysis revealed that acquisition differentials between the consistent and inconsistent blocks were significantly different based on the order (z = -2.597, p = .009, two-tailed; large effect size), with differentials typically larger for those who received the inconsistent block first (M = 21.40) relative to the consistent block first (M = 12.55). However, it is important to note here that there appears to have been an unfortunate proliferation of one order over the order, generated by chance through the use of a randomised block order feature of the software employed, rather than by conscious

counter balancing. This imbalance in orders may partly account for the absence of an order effect as ascertained through inductive statistical analyses.

Table 2.2.6: Running order for inconsistent and consistent test blocks (1: Consistent first, 2: Inconsistent first)

Group One															
Participant	1	5	7	8	9	10	11	12	13	14	15	16	18	21	30
Order	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2
Group Two															
Participant	2	3	4	6	17	19	20	22	23	24	25	26	27	28	29
Order	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2

2.2.3.6 Summary

This experiment again revealed a significant reduction in baseline acquisition rates across the two baseline blocks for all participants, again indicating a significant and worrying practice effect. With regard to the two FAST testing blocks, the trials acquisition rates for the inconsistent and consistent blocks were not significantly different to one another for Group One, suggesting no overall FAST effect for this group. However, there was a significant difference in trial requirements between the two FAST blocks for Group Two, albeit in the unexpected direction. That is, participants with high MHS actually required significantly more trials to achieve fluency for the inconsistent block relative to the consistent block, suggesting a strong pre-existing Gay-Good relation for this sample. SoR distribution reflected a similar trend, with Group Two reporting significantly higher Gay-Good SoR indices than Group Two, in addition to reporting a statistically significant overall Homopositivity FAST effect. This would suggest, therefore, a weak relationship between the FAST and self-reported homonegativity. Finally, this experiment revealed that, in contrast to Experiment 1, the order in which the two testing blocks were administered did significantly affect acquisition rates for the two FAST blocks.

2.2.4 Discussion

The aim of the current experiment was to assess the suitability of multi-exemplar pictorial stimuli in a Homopositivity FAST, which was assessed largely via what is known as a 'known-groups' approach. It was hypothesised that, based on their verbal histories, FAST scores for Group Two (moderate-to-high explicit homonegativity) would be notably different to those in Group One (low explicit homonegativity). More specifically, it was expected that individuals in Group Two would exhibit behavioural resistance to the formation of functional Gay-Good stimulus classes, as represented by pictorial exemplars. In contrast, a more varied outcome was expected for Group One, given that lower MHS scores might be more prone to social desirability distortions than higher scores. Additionally, as in Experiment 1, it was conjectured from previous research that there would be a general trend in the data toward weak or reversed Gay-Good SoR indices (i.e., that there would be an overall FAST effect), given the overall anti-gay bias that would be expected in a typical Irish population (e.g., O'Higgins-Norman et al., 2006).

However, following statistical analyses, it would appear the current experiment showed none of the predicted effects. Firstly, there was no significant overall bias in either direction observed in the current sample, which was unexpected, with FAST scores actually tending toward positive values for the sample as a whole. Furthermore, while Group One showed expected variance in scores with no particular bias in either direction, Group Two actually exhibited a significant pro-gay bias (i.e., a strong FAST effect), which would not have been expected based on their high levels of self-reported homonegativity. In addition to this, when FAST scores from the two groups were compared, Group Two actually showed significantly more positive SoR indices than Group One. Thus, it can be inferred from the outcomes here that the current stimuli are also not effective exemplars in a Homopositivity

FAST, given that it failed to engender any expected differences in learning rates across test blocks through a known-groups paradigm.

Though this outcome was unexpected, there are some possible reasons why this may have been the case. Firstly, as previously discussed, it is of considerable importance to ensure that the exemplars selected depict the target concepts as closely as possible in order to make participants respond to the label in the desired way (e.g., De Houwer, 2002; Govan & Williams, 2004; Olson and Fazio, 2003). That is, exemplars should be evaluated solely on the basis of their membership in the appropriate class (Steffens & Plewe, 2001; Lane et al., 2007). However, while the men chosen for the images are widely known Gay celebrities, it is possible that images were categorised as belonging to a likeable subtype (well known Gay celebrities), resulting in confounding cross-class evaluations (see Bluemke & Friese, 2005; De Houwer, 2001). In this instance, the individual exemplars may have contained positive stimulus functions and, as such, may not have represented participants' evaluations of the category as a whole (e.g., 'nice' Gay people rather than Gay people altogether). Therefore, even in the case where a participant may generally dislike Gay people, the activation of exemplars eliciting strong positive functions may have led to a temporarily different evaluation of the category. This may explain the tendency in the current sample toward positive SoR values, given that it could explain the significant pro-gay bias observed even for strongly anti-gay individuals. Specifically, it may be the case that in the context of picture stimuli, different histories of stimulus-stimulus relations are measured than are measured using a word-based MHS questionnaire. This is merely conjecture at this point and there is no easy way to resolve this issue except empirically due to the fact that some known outcome is required in order to validate any new test. In time and across experiments it may emerge, however, that word and picture based "implicit" tests can sometimes measure different relations present in the history of the participant, and that these different relations may not

even cohere with each other (e.g., the word ‘Gay’ may be negative but an image of a Gay person may be positive for a given individual).

In a similar manner, a lack of familiarity with the celebrities employed here may have also influenced FAST performance. De Houwer (2001) reiterated the importance of eliminating irrelevant features of stimuli that may cause participants to categorise them in a way that is unintended (i.e., in this case other than as a Gay individual). While the celebrities used here would be widely known Gay public figures, ambiguity or confusion surrounding their sexual orientation of may still have been possible. In this instance, the most salient feature of the images may not have been the sexual orientation of the individuals, but rather additional confounding and extraneous features of the images, such as their fame and general public popularity. In this way, it may have made it possible that exemplars were categorised based on features that were irrelevant to the current construct (e.g., ‘White’, ‘Male’ or ‘Celebrity’ rather than ‘Gay’, etc.). Considering again, for example, the study by Mitchell and colleagues in which implicit preference for White people over Black people was diminished significantly when positive exemplars are used to represent Black individuals and negative exemplars for White individuals (Mitchell et al., 2003). In this case, the current measure may have been assessing pro-man or pro-celebrity bias, rather than investigating the intended Gay-Good stimulus relation, which may have reversed the effect. Consequently, it is possible that either cross-class associations (De Houwer, 2001; Bluemke & Friese, 2005) or indeed a lack of familiarity with the sexual orientation of the celebrities could have resulted in the outcomes observed here. That is, in the event of such stimulus ambiguity, observed outcomes may be unrelated to MHS scores, as was the case here.

One way to eliminate the foregoing confounding features would be to shift the focus of images away from public figures and utilise more generic images of Gay individuals. However, it is difficult to convey conclusively that an individual is Gay using exemplars with

one solitary person. Thus, it may more effective to select exemplars that depict same-sex couples, rather than Gay individuals. This technique has been utilised previously in IAT research looking at implicit attitudes toward Gay versus straight individuals, where images of couples rather than individuals are employed as exemplars of the target constructs (e.g., Banse et al., 2001; Jones & Devos, 2013; Jost et al, 2004; Rohner & Björklund, 2006). While the current experiment is more interested in obtaining an absolute level of prejudice toward Gay individuals exclusively rather than a level of preference for opposite versus same-sex couples, using couple exemplars may be a way to circumvent the ambiguity of the individuals' sexual orientations in future experiments.

It may also be necessary to further enhance the affective valence of the stimuli, given that an additional reason the current exemplars were ineffective could have been that they were still not evocative enough. As discussed previously, research suggests that anti-gay prejudice has a strong affective component (e.g., Hudepohl et al., 2010; Kimmel, 1997; Mahaffey et al., 2005; Parrott, 2008). While the pictorial exemplars used here were selected based on their perceived strong visceral or evocative response functions relative to the previous verbal stimuli, it is possible that they still served as purely semantic representations of the construct (e.g., 'a Gay male'), rather than of the construct as a whole (e.g., same-sex romantic or sexual behaviour). In a similar vein, it may be possible that the exemplars chosen to represent the 'Good' category (positive images of people skiing) may have been ineffective, given that their face validity as strongly positive and evocative stimuli is somewhat unclear. Thus, in future experiments, it may be beneficial to further enhance the affective valence of the exemplars selected to represent both of the target constructs, in order to ensure the stimuli are more suited for exemplifying functional stimulus classes of emotive stimuli (i.e., 'Gay' and 'Good').

An additional issue that needs to be addressed here is the order effects that were observed in the current data set. The analyses conducted above revealed that, in contrast with Experiment 1, the order in which the two FAST blocks were administered (i.e., whether the inconsistent or consistent block was presented first) impacted significantly upon test results. While it is difficult to interpret this finding in light of the other issues in the experimental design, its potential influence upon the outcomes should nonetheless be discussed. Previous research suggests that the order in which IAT blocks are administered matters, with studies suggesting that IAT effects tend to be larger in instances where the compatible phase precedes the incompatible phase, as was the case here, rather than vice versa (see Klauer & Mierke, 2005; Nosek et al., 2003; Lane et al., 2007, for reviews). Consequently, it is widely recommended in the IAT literature to counter-balance the order of the test blocks so this effect is minimised at a group level. However, in the case of Experiment 2, while the order was indeed randomised, there was unfortunate proliferation of one order over the other, with approximately two thirds of participants receiving the consistent block first. This may not only have enhanced order effects across the group as a whole but the low number of participants exposed to other block sequence may have been so small as to render all statistical inference regarding order effects unreliable. Investigating this issue with dedicated experiments is beyond the scope of the current research. However, it is an issue to be borne in mind, and future experiments should ensure a more balanced randomization of the block order. Similarly, with regard to the observed practice effects, it is again the case that dedicated analyses would represent a distraction from the main purpose of creating an in-principle Homopositivity FAST prototype. However, this issue will be observed in subsequent experiments and suggestions will be made for its amelioration in future research.

In conclusion, the current experiment documented none of the predicted effects, with an unexpected pattern observed in the data set. Consequently, it may be logical to infer here

that the stimuli were unsuccessful in accurately representing the target constructs (i.e., assuming the core process is functional as reported by O'Reilly e al., 2012). This may have been the result of confounding intra or cross-category associations, given that the exemplars chosen may have been evaluated as members of a likeable subtype (i.e., 'Gay celebrities') rather than as members of the overall stimulus category ('Gay'). Additionally, it may have been a lack of familiarity with the individuals depicted in the Gay stimuli, or of their sexual orientation, that affected outcomes. Alternatively, it may have been that the particular exemplars employed were still not evocative enough to engender measurable behavioural resistance in class formation. Consequently, the following chapter will focus on the relative effectiveness of pictorial stimuli with enhanced affective valence in a Homopositivity FAST.

Chapter 3

Enhancing the Affective Valence of Pictorial Stimuli: Effects on FAST Outcomes

Experiments 3a and 3b

3.1 Experiment 3a

3.1.1 Introduction

As previously discussed, one of the most important elements in development of any new implicit measure is establishing the most appropriate stimulus exemplars that will be employed. Thus far, this research programme has been concerned with identifying whether pictorial or verbal exemplars are the most suitable for a Homopositivity FAST. The findings from these experiments in conjunction with previous research appeared to suggest that pictorial exemplars would be more appropriate for an affective construct such as homonegative prejudice. More specifically, what the previous experiments appeared to indicate was the importance of selecting pictorial exemplars that are both unambiguous and high in affective valence in order to evoke the emotional stimulus functions associated with the stimulus class (i.e., Gay-Good).

As stated above in the discussion section of Experiment 2, one of reasons the images may have been unsuitable exemplars of the category ‘Gay’ could have been due to ambiguity caused by the celebrity status of the individuals used. While this celebrity status would have ensured that the majority of participants would have been familiar with the men in the images, it may have been problematic if participants were unaware of sexual orientation of the celebrities used, or if celebrity status was the defining feature of the class of exemplars (i.e., rather than sexual orientation). Even if the sexual orientation of all exemplars was salient, it was surely confounded by the positive functions elicited by celebrities for many people. In other words, even if an individual is generally prejudicial towards Gay people, this prejudice may not manifest for liked popular publicly Gay figures. Research has suggested that likeable exemplars (such as well-liked Gay celebrities) may diminish test effects. As previously discussed in Chapter 2, it has been suggested that the context in which the exemplars are presented may exert significant influence over how participants categorise and

engage with the super ordinate stimulus category (Lane et al., 2007). Considering again, for instance, the study by Mitchell and colleagues (2003), which revealed that when disliked Black people and liked White people represented the categories Black and White, participants typically exhibited significant pro-White bias, relative to when liked Black people and disliked White people, which was associated with significantly diminished pro-White preference (Mitchell et al., 2003). In this way, the images used in Experiment 2 may have been positively valenced, rather than being negatively valenced as intended, which can reduce implicit prejudice towards exemplars (Chiu, Ambady & Deldin, 2004; Govan & Williams, 2004).

Thus, Experiment 3a of the current thesis will gain greater control over the target stimulus category by ensuring that the homosexual content in the images is more salient, and ideally the most salient feature of the images. This will be achieved by employing images of unknown men (i.e., not widely known public figures) to ensure that the target construct is appropriately (and unambiguously) depicted. Additionally, in a similar manner to previous IAT research (e.g., Jones & Devos, 2013; Jost et al, 2004; Rohner & Björklund, 2006), the current experiment will employ images of same-sex couples rather than solitary individuals, which should hopefully eliminate any confusion surrounding the orientation of the individuals.

As previously discussed, it would appear to be a compelling case that homonegative attitudes may be characterised by strong emotional, rather than purely "verbal" responses, and thus emotionally evocative stimuli should be employed as exemplars (Hudepohl et al., 2010; Kimmel, 1997; Mahaffey et al., 2005; Parrott, 2008). However, as discussed above, an additional issue with the pictorial exemplars employed in Experiment 2 could have been that the affective element of the images was not salient enough. In other words, the images of solitary individuals were unable to instantiate the construct (i.e., 'Gay') in the desired way.

Therefore, in addition to minimising the ambiguity of the images, it may also be necessary to further enhance the affective valence of the pictorial exemplars through depicting a more explicit same-sex relationship rather than an individual.

Thus, it is of considerable importance to establish the precise way in which the affective valence of the images should be enhanced in order for the exemplars to be representative of the wider construct (i.e., same-sex relationships). One possible way to improve upon the representativeness of the pictorial exemplars could be to enhance the salience of the group membership of the individuals in the images. Previous prejudice research examining attitudes towards African-Americans in the US has suggested that the degree to which prejudiced behaviour is manifested may depend not just upon group membership (e.g., an image of a Black male), but also the extent to which the target possesses other stereotypical group-associated qualities (e.g., aggressive or violent expressions), such that the salience of the individual's group membership may potentially activate more powerful associated stereotypes associated with that group (Fiske, Cuddy, Glick, & Xu, 2002). Chiu and colleagues suggested, therefore, that the valence of group-associated emotion expressed by an individual group member is likely to affect the response of a perceiver to the targets, given that it could potentially enhance the activation of group stereotypes (Chiu et al., 2004).

Therefore, it is necessary to establish the most salient affective feature associated with Gay men and ensure that this is the primary focus of the exemplars. It has been suggested that intimate behaviour between two men is a particularly pronounced feature associated with homonegative prejudice (see Kimmel, 1997). Indeed, previous research examining homonegativity has revealed that man-man intimacy is associated strongly with negative attitudes toward Gay men (Pettigrew & Meertens, 1995; Lacerda, Pereira & Camino, 2002). One study, for example, revealed that images depicting both erotic (i.e., sexual intimacy

between two men) and non-erotic intimacy (i.e., normal relationship behaviours such as holding hands, kissing, dating, etc.) were strongly associated with anti-gay aggression among homonegative individuals (Hudepohl et al., 2010). Thus, the following experiment will attempt to enhance the affective valence of the Gay images exemplars by incorporating elements of romantic intimacy between the Gay men, (e.g., by using images of men kissing, cuddling, and getting married).

Experiment 1 and 2 suggested that the stimuli utilised in those experiments were not appropriate exemplars for the target 'Gay' stimulus category. As just discussed, it is of considerable importance to ensure that the category 'Gay' is adequately represented both in terms of the emotional stimulus functions of Gay stimuli as well as the semantic ones. However, it is equally crucial to re-consider the positively valenced images used for the 'Good' category as well. As discussed previously, research suggests that anti-gay attitudes are evaluative in nature. Consequently, it may be wise to explore the use of more positively valenced 'Good' category images. Experiment 2 employed images of people skiing, which were rated by the IAPS database as high in positive valence and arousal. However, it is possible that the images were higher in positive valence only on the semantic rating scale, based on which the IAPS valence ratings were first collected. Although the IAPS manual also ranks the skiing images as high in arousal, it is unclear what type of arousal this refers to (i.e., positive or negative, or even varying across individuals). In other words, while skiing images may be rated semantically as 'good', their face validity as evocative stimuli is somewhat ambiguous. To investigate this, a more common-sense approach will be adopted here for choosing images on the basis of face validity, rather than relying on IAPS ratings alone. The current experiment, therefore, will utilise images of young animals (e.g., a group of kittens) with similarly high IAPS ratings for arousal and valence, but which are also more viscerally positive and salient as emotionally positive stimuli.

The current experiment will also be modified relative to the two previous experiments in order to incorporate a 'neutral FAST' into the procedure to function as a sort of baseline against which to interpret main FAST effects. Specifically, given that the current experiment was assessing the relational strength of the pre-existing Gay-Good relations for participants, it was conjectured that it might be appropriate to include a block that is conceptually identical to the Gay-Good block, but that differed only in terms of the romantic content of the images. This approach is quite common within the implicit testing literature for comparison purposes, and has been used in a variety of contexts. These have included, for example, implicit alcohol associations in those with alcohol addiction, whereby an alcohol-neutral block was included alongside an alcohol-positive and alcohol-negative IAT (Dickson, Gately & Field, 2013; Jajodia & Earleywine, 2003). These neutral test blocks have been used in a similar way within other contexts, such as for implicit gambling (Brevers, Cleeremans, Hermant, Tibboel, Kornreich, Verbanck & Noel, 2013) and implicit cannabis associations (Dekker et al., 2009). The current research aims to use a similar approach by including a 'pro-men' (i.e., as opposed to a 'pro-gay') FAST into the current procedure. This neutral FAST will be identical in every respect, save for the substitution of the Gay images with matched pictures of pairs of men in non-romantic and sexual situations (e.g., shaking hands). In effect, the primary difference between the two FAST exposures will be the romantic intimacy between the two men, which should allow for a clearer test outcome, insofar as it will facilitate a clear comparison between a pro-man and pro-gay stimulus relation strengths. Thus, what is being aimed for is not so much to detect a straightforward pro-gay FAST effect, but to identify a difference in the pro-gay compared to pro-man bias across the two FAST exposures.

In summary, Experiment 3 aims to assess the effectiveness of more emotive and unambiguous pictorial stimuli in a Homopositivity FAST, and to obtain in-principle proof that the general FAST procedure is viable and can be improved as a tool for assessing attitudes

towards sexual minorities. This experiment will again adopt a known-groups approach, wherein the MHS will be used to establish self-reported levels of explicit homonegativity against which FAST scores can be compared. This experiment additionally aims to investigate the usefulness of including a neutral pro-man FAST alongside the pro-gay FAST for comparison purposes.

It is expected that there will be some general correlations between scores on the MHS and the two-phase Homopositivity FAST. More specifically, it is expected that increases in MHS scores will be generally correlated with greater behavioural resistance for the pro-gay FAST (i.e., weaker and more negative SoR indices) and with greater behavioural facilitation for the pro-men FAST (i.e., stronger and more positive SoRs).

3.1.2 Method

3.1.2.1 Participants

Participants for this experiment were recruited from either the NUI, Maynooth undergraduate population or from acquaintances of the experimenter via word of mouth recruitment. All participants were adults (over 18 years of age) and did not suffer from any condition that would interfere with their ability to perform routine cognitive tasks. The sample (n = 5) was comprised of 2 males and 3 females aged between 23 and 28 (M = 24.4; SD = 2.19), who were all either attending University or who were recent University graduates. The low sample reflects the fact that the experiment was abandoned due to emerging methodological issues that will be addressed in the Results and Discussion section, as well as in Experiment 3b.

3.1.2.2 Ethical Considerations

As discussed previously, all elements of this research process received prior approval from the NUI, Maynooth Ethical Committee, and were all in accordance with the

Psychological Society of Ireland (PSI) code of ethics, and with the NUI, Maynooth Department of Psychology guidelines for safe working practice in psychological research.

3.1.2.3 Materials

3.1.2.3.1 The FAST Software Package and Stimulus Set

The version of the FAST utilised in the current experiment was again developed with the computer programming software Livecode™, and presented on a 13" laptop (1280x800 pixel resolution). Image stimuli for this experiment were again predominantly obtained from the IAPS database (Lang, et al., 1999). This is with the exception of the images representing the categories 'Gay' and the neutral FAST category 'Men Shaking Hands', which were purchased from an online stock photo database (www.istockphoto.com), in addition to a small selection of stimuli used during the practice block. Stimuli obtained from the IAPS were matched roughly for standardised arousal ratings and valence across all blocks except for the consistent and inconsistent blocks during the testing portions of the FAST, in which one of the stimulus categories consisted of positively valenced and highly arousing images (as rated by the IAPS manual). Images were cropped slightly to achieve a perfectly square shape and in a small number of cases to eliminate peripheral objects in the image view. All images were 150x150 pixels in size (at a resolution of 72 dpi). During all blocks, stimulus categories were comprised of four image exemplars. The alphanumeric category code, category description, and their corresponding IAPS/istockphoto image identifier numbers are listed below (Table 3.1.1).

Table 3.1.1: Pictorial exemplars employed in Experiment 3a

Alphanumeric	Stimulus Category	IAPS Image Number/Image Description/istockphoto Image Number
Phase 1: Practice block		
P1	Circles	Four discriminable images of circles in varying sizes and colours

P2	Squares	Four discriminable images of squares in varying sizes and colours
Q1	Chinese Characters	Four discriminable Chinese-type symbols
Q2	Numbers	Numbers 1, 2, 3 and 4 in black 48-type font
Phase 2: Baseline 1		
X1	Mushrooms	5531, 5532, 5533, 5544
X2	Cakes	7200, 7250, 7270, 7282
Y1	Birds	1333, 1419, 1450, 1740
Y2	Furniture	7025, 7235, 7705, 7710
Phase 3: FAST [1]		
A1	Small animals	1440, 1460, 1710, 1750
B1	Gay men	1917875, 2007866, 3657240, 10197611
N1	Nature	5201, 5220, 5760, 5780
N2	Vehicles	7130, 7140, 7595, 8510
Phase 4: FAST [2]		
A1	Men shaking hands	728463, 14316678, 5177840, 7187079
B1	Small animals	1440, 1460, 1710, 1750
N1	Nature	5201, 5220, 5760, 5780
N2	Vehicles	7130, 7140, 7595, 8510
Phase 5: Baseline 2		
X3	Sky	5593, 5594, 5990, 5991
X4	Babies	2040, 2050, 2250, 2260
Y3	Food	7281, 7291, 7350, 7480
Y4	Zoo Animals	1601, 1661, 1810, 1812

3.1.2.3.2 The Modern Homonegativity Scale.

The current experiment also utilised an identical 10-item MHS scale for the purposes of identifying levels of self-reported homonegativity among participants.

3.1.2.4 Design

While it was intended to employ an identical group-based design as in previous experiments, given that the experiment was abandoned in its early stages, a simple correlational design was conducted wherein the scores on the MHS will be correlated with the SoR values from both FAST blocks. That is, the purpose of the analysis was to test the hypothesis that higher explicit homonegativity is correlated with weaker or more negative SoRs for the Gay-Good test bock relative to the Men-Good test block.

3.1.2.5 General Experimental Sequence

3.1.2.5.1 Informed Consent

Informed consent was obtained from all participants using identical instructions as in previous experiments.

3.1.2.5.2 Function Acquisition Speed Test

Upon supplying written consent, the participant was then required to complete the FAST portion of the experiment. This involved firstly ensuring the participant was comfortable, the environment was quiet, and that there were limited external distractions. The participant was then seated at a desk and presented with the FAST programme on a laptop computer. The FAST procedure adhered to the same general experimental format outlined in Experiments 1 and 2, with the only difference being that two testing phases were employed rather than one (i.e., a practice phase, two baseline phases, and *two* testing phases). The reinforcement contingencies for each phase of the FAST are outlined below.

Phase 1: Practice block

The first phase of the FAST procedure was the practice block, during which the stimulus categories P1 (circles) and P2 (squares) shared the response key 'Z', while categories Q1 (Chinese characters roughly translating as 'nonsense') and Q2 (numbers) shared the response key 'M'. Participants were presented with identical instructions to experiments one and two.

Phase 2 and 5: Baseline blocks

During the first baseline block, the stimulus categories X1 (mushrooms) and X2 (cakes) shared the response key 'Z', while categories Y1 (birds) and Y2 (furniture) shared the response key 'M'. During the second baseline block, the stimulus categories X3 (sky) and X4 (babies) shared the response key 'Z', while categories Y3 (food) and Y4 (zoo animals) shared the response key 'M'. Both baseline blocks continued presenting stimuli in a randomised order until participants reached fluency or until they reached 200 trials.

Phases 3 and 4: Testing blocks

Both Phase 3 and Phase 4 consisted of a consistent block and an inconsistent block, which were presented in a random order. Specifically, within the Gay-Good block, during the inconsistent block, the stimulus categories A1 (small animals) and N1 (nature) shared the response key Z, while categories B1 (Gay men) and N2 (vehicles) shared the response key M. In the consistent block, the stimulus categories A1 (Gay men) and B1 (small animals) shared the response key Z, while categories N1 (nature) and N2 (vehicles) shared the response key M. Within the Men-Good block, during the inconsistent block, the stimulus categories A1 (men shaking hands) and N1 (nature) shared the response key Z, while categories B1 (small animals) and N2 (vehicles) shared the response key M. In the consistent block, the stimulus categories A1 (men shaking hands) and B1 (small animals) shared the response key Z, while categories N1 (nature) and N2 (vehicles) shared the response key M. It is important to note here that the two target categories, A1 and B1, were not identical across the two FAST

exposures. That is, the 'small animals' category was not always B1 and the target (i.e., the Gay men or the men shaking hands) was not always A1. This was merely to eliminate any potential confounds from an identical format, which could make the first FAST comparatively easier than the first regardless of the individual's verbal histories.

3.1.2.5.3 The Modern Homonegativity Scale

The MHS was again administered following the FAST portion of the experiment as a means of assessing participants' levels of self-reported homonegativity.

3.1.2.5.4 Debriefing

Participants were fully debriefed regarding the nature of the experiment immediately upon completion in an identical manner to previous experiments. Participants were also asked at this stage if there was any ambiguity or confusion surrounding the experimental stimulus categories used in the FAST.

3.1.3 Results and Discussion

This current experiment was abandoned in its early stages, and consequently the following results discussed here consist of merely preliminary analyses of the data set. The reasons for abandoning this experiment and revising it as Experiment 3b will be outlined below.

3.1.3.1 Modern Homonegativity Scale Scores

With possible scores on the MHS ranging from 10-50, it can be ascertained from the scores below that the sample are relatively low in explicit homonegativity (M = 15.2, SD = 4.66), which is not unexpected given the explicit pro-gay bias that may be more prevalent in an educated University population (see Cerny & Polyson, 1984). That is, research has suggested that people with a higher level of education tend to report lower levels of homonegative prejudice generally (e.g., Elchardus & Spruyt, 2009; Ohlander, Batalova, &

Treas, 2005), and, as such, the current sample may not be as homonegative as a more typical Irish sample.

Table 3.1.2: MHS Scores for all participants

Participant	MHS
1	22
2	15
3	10
4	17
5	12

3.1.3.2 Practice Blocks

As was the case for previous experiments, practice block data will not be discussed here.

3.1.3.3 Baseline Blocks

The number of trials to criterion for each participant across both baseline blocks is represented in Table 3.1.3 below. An inspection of the data set reveals that, as was the case in previous experiments, acquisition rates for the second baseline block are notably lower than in the first. Indeed, as with previous experiments, a Wilcoxin Signed Ranks test revealed a statistically significant difference between the number of trials to fluency at the first and second baselines for the entire sample (z = -2.023, p = .043, two-tailed).

Table 3.1.3: Number of trials to fluency criterion across both baseline blocks (Phases 2 and 5)

Participant	Baseline 1	Baseline 2	Mean Baseline
1	24	12	18
2	27	12	19.5
3	14	11	12.5
4	59	22	40.5
5	25	16	20.5

3.1.3.4 Function Acquisition Speed Test

3.1.3.4.1 Acquisition Rates

There were two test blocks in the current experiment, Phase 3 and Phase 4, which were presented in a randomized order. Table 3.1.4 documents the number of trials to criterion for the Gay-Good FAST block, while Table 3.1.5 documents the number of trials to criterion

for the Men-Good FAST block. Acquisition rates are generally in the direction that is not unexpected for explicitly pro-gay individuals, with FAST trial requirements revealing that participants typically found the Gay-Good block easier (in terms of the number of trials to criterion) than the Men-Good block, indicating a pro-gay bias in the current sample. A Wilcoxin Signed Ranks test revealed that participants required significantly more trials to criterion for the inconsistent relative to the consistent block for the Gay-Good block (z = -2.23, p = 0.043, two-tailed), suggesting a strong FAST effect at the group level. In contrast, for the Men-Good FAST block, a Wilcoxin Signed Ranks test revealed no statistically significant difference in acquisition rates between the inconsistent and consistent blocks (z = -1.095, p = .273, two-tailed), suggesting a non-existent FAST effect (i.e., a weak or non-existent men-good relation). Of course, it should be reiterated that the sample size employed here is very small and intended only to give general indication of the performances of participants using the current procedure.

Table 3.1.4: Number of trials to criterion for the Gay-Good FAST block (Phase ¾. Note: the difference is obtained by subtracting the number of trials to criterion for the consistent block from the inconsistent block.

Participant	Inconsistent	Consistent	Difference
1	25	24	1
2	23	11	11
3	20	11	9
4	17	14	3
5	33	22	11

Table 3.1.5: Number of trials to criterion for the Men-Good FAST block (Phase 3/4)

Participant	Inconsistent	Consistent	Difference
1	17	17	0
2	39	27	12
3	14	12	2
4	23	26	-3
5	17	11	6

3.1.3.4.2 Strength of Relation Indices

Table 3.1.6 documents the SoR Indices for both the Gay-Good and Men-Good FAST blocks, which were calculated using the same formula as previous experiments. The current study expected that, for pro-gay individuals, there would be a general trend toward positive SoR indices for both FASTs. Single-sample t-tests revealed a strong overall effect for SoR indices for the pro-gay FAST t (4) = 2.947, p = .041, two-tailed, though a non-significant effect for the pro-men FAST, t (4) = 1.568, p = .192, two-tailed. Indeed, as evidenced by the table below, participants typically report stronger, more positive SoRs (indicative of a stronger pre-existing relation) for the Gay-Good block relative to the Men-Good block, though a Wilcoxin Signed Ranks test revealed this to be non-significant (z = -1.826, p = .068, two-tailed).

Table 3.1.6: Strength of Relation [SoR] Indices for both FAST blocks. Note: large positive SoRs indicate a strong pre-existing relation, while zero or negative SoRs suggest a weak or absent pre-existing relation

Gay-Good FAST block		**Men-Good FAST block**	
Participant	[SoR]	**Participant**	[SoR]
1	0.79	**1**	0
2	9.3	**2**	9.3
3	8.18	**3**	1.82
4	1.86	**4**	-1.86
5	8.39	**5**	4.60

3.1.3.4.3 Order Effects

It was also necessary to investigate any potential differences in the data set based on the order in which the inconsistent and consistent blocks were administered. Tables 3.1.7 and 3.1.8 below document the order in which the inconsistent and consistent blocks were administered for the Gay-Good and Men-Good FASTs, respectively. To investigate any order effects, Mann-Whitney U tests were conducted comparing trial acquisition differentials between the consistent and inconsistent blocks for both FAST blocks. These analyses revealed that, within both the Gay-Good block and the Men-Good block, trial differentials

between the inconsistent and consistent block were not significantly different based on order (p = 7.67; p = . 83, respectively).

Table 3.1.7: Running order for inconsistent and consistent test blocks (1: Consistent first, 2: Inconsistent first) for the Gay-Good *FAST*

Participant	1	2	3	4	5
Order	2	1	1	1	1

Table 3.1.8: Running order for inconsistent and consistent test blocks (1: Consistent first, 2: Inconsistent first) for the Men-Good FAST

Participant	1	2	3	4	5
Order	1	2	2	1	2

Similarly, it was important to assess whether FAST scores were influenced by whether the Gay-Good or Men-Good FAST was administered first. Table 3.1.9 below documents the order that the two FASTs were administered for each participant. To investigate any order effects, it was necessary to compare SoR indices for each participant's first block with their second block. An inspection of the order alongside the SoR indices (Table 3.1.6) does not suggest any particular trend based on the order.

Table 3.1.9: Running order for FAST blocks 1: Men-Good first, 2: Gay-Good *first*

Participant	1	2	3	4	5
Order	1	2	1	1	2

3.1.3.6 Summary

The aim of the current experiment was to investigate the effectiveness of multi-exemplar pictorial stimuli, which were high in affective valence, in a two-phase Homopositivity FAST. Similar to previous experiments, it was hypothesised that there may be general patterns visible within the data based on MHS scores. Based on this current sample's low-to-moderate levels of explicit homonegativity, therefore, a general overall trend toward stronger and more positive SoR indices for the pro-gay FAST relative to the pro-men FAST was expected. Though the very small sample size should be taken into account when interpreting analyses, a preliminary statistical investigation appeared to support this hypothesis, revealing a strong overall pro-gay bias in the sample. Firstly, in terms of

acquisition rates between the consistent and inconsistent blocks, participants required significantly fewer trials to reach criterion in the consistent block relative to inconsistent block for the pro-gay block, which would be indicative of a strong pre-existing Gay-Good stimulus relation. In contrast, no significant different was found between these two blocks within the pro-men FAST. Similarly, there was a significant overall effect for SoR indices for the pro-gay FAST, which was absent for the pro-men FAST. Therefore, based on this first preliminary administration to five participants, it would appear that the current Homopositivity FAST employing affective pictorial stimuli is sensitive to the expected pro-gay bias in an educated University population.

Additionally, these analyses revealed no significant order effects within the current data set, for either the order of the test blocks within each FAST (i.e., the order in which the consistent and inconsistent blocks were administered) or between the two FASTs (i.e., whether the pro-gay FAST or pro-men FAST was administered first). However, these preliminary analyses again revealed a significant reduction in baseline acquisition rates across the two baseline blocks for all participants, with participants improving substantially across each phase of the FAST. In conjunction with the previous experiments, this provides further evidence for a worrying practice effect within the FAST procedure. However, as mentioned previously, this will remain a simple observation until the more pressing issues with stimuli selection have been sufficiently addressed.

The reason for the low sample size in the current experiment relates to the fact that the current experiment relied upon finding anti-gay individuals as participants (i.e., a known groups paradigm) in order to assess the effectiveness of the FAST. The MHS scores can be used for this purpose; however, the number of individuals in the population displaying notable anti-gay bias according to the MHS is relatively low. Given that large samples are needed to validate a new implicit test of this kind, obtaining substantial samples of anti-gay

individuals may not be feasible considering the nascent nature of the procedure. That is, it would be necessary to secure high numbers of anti-gay individuals in order to validate the current procedure, whereas a procedure that could be validated against a more varied or even a largely explicitly pro-gay sample would be more pragmatically useful.

It was decided, therefore, that while the current data appears to suggest that the current procedure is broadly viable, a procedure that does not reply upon prior anti-gay bias for validation would be beneficial (i.e., a paradigm within which even a homogenous or more explicitly pro-gay group could be employed). That is, it was realised that a test that could compare an individual's positive and negative biases towards homosexual individuals would be a more powerful test, and allow any population sample to be employed for validation because (a) anti or pro-gay bias would be easier to detect within participants rather than across and (b) scores on two 'opposing' FAST tests should be different to each other (that is, they should not positively correlate), thereby constituting a further check on the validity of the measure. The use of two FASTs for both positive and negative bias, therefore, might be likened to the use of lie scales within questionnaires, in which a questions is asked several times in different semantic directions (e.g., "are apples nicer than oranges?" and "are oranges nicer than apples"?) in order to establish a level of consistency in participant responses. Addressing this issue was the purpose of Experiment 3b.

Experiment 3b

3.2.1 Introduction

The previous experiment provided in-principle support for the general format of a Homopositivity FAST using the pictorial stimuli outlined in 3a. However, as discussed in the Results and Discussion of Experiment 3a, it was decided after administering the test to five participants that a FAST for anti-gay attitudes would benefit from removing the neutral 'control' FAST (i.e., the Men-Good FAST) and replacing it with a Homonegativity (i.e., a Gay-Bad) FAST. The current experiment, therefore, is almost identical to Experiment 3a. However, the neutral FAST will be replaced with a FAST measuring a contrasting Gay-Bad affective stimulus class, in which the same Gay images as in the Gay-Good FAST will be employed, but negatively valenced target images will be used in the place of the positively valenced small animal pictures.

A considerable challenge for this experiment, therefore, lay in identifying stimuli that could accurately instantiate the concept 'Bad' as it best pertains to anti-gay attitudes. Naturally, as with any affective construct, anti-gay attitudes are complex and varied, and so could potentially vary substantially across participants. However, research has suggested that disgust is a particularly common primary emotional response toward Gay individuals (Cottrell & Neuberg, 2005; Haidt, Rozin, McCauley, & Imada, 1997; Olatunji, 2008), having been found to related to both implicit and explicit sexual prejudice (Inbar, Pizarro, Knobe, & Bloom, 2009; Hodson & Costello, 2007). Consequently, the current research will use 'disgust' images as the particular variety of negatively valenced target images in the Homonegativity FAST. In effect, this novel FAST will measure the strength of a pre-existing Gay-Disgust relation for participants, alongside a homopositive or Gay-Good relation, as measured in the previous experiment.

It is again expected that there will be some general correlations between scores on the MHS and the two-phase Homopositivity/Homonegativity FAST. More specifically, it is expected that increases in MHS scores will be generally correlated with greater behavioural resistance to the formation of a Gay-Good stimulus class relative and with greater behavioural facilitation to the formation of a Gay-Disgusting stimulus class. Thus, it is expected that there will be an overall difference between the two sets of SoR indices for participants that should reflect MHS scores. However, it is important to note that the two FAST scores are not necessarily expected to mirror each other (i.e., whereby an SoR of +5 would be associated with a Homonegativity SoR of -5). Rather, as discussed above, the inclusion of a second Homonegativity FAST serves to stabilise the findings from the Homopositivity FAST by examining the strength of the opposite stimulus relation. That is, should an anti-gay individual exhibit strong resistance to the formation of a Gay-Good stimulus class, it would not be coherent or logical if they exhibited the same resistance to the formation of a Gay-Disgusting stimulus class. In other words, it is hypothesised that these two sets of scores should not positively correlate.

In addition to this, it is important to clarify it is not expected that there will be an overall bias in any direction given that the sample will be recruited from a University-educated population. Indeed, it should be noted that this is not the aim of the study, and that an overall anti-gay group effect is not required in order to validate the FAST using the current design. Rather, the FAST will be assessed in terms of its ability to provide two distinct scores for each of two FASTs that assess inverse attitude biases (i.e., pro and anti-gay). In addition, using this ‘double FAST’ method it will be possible to calculate an overall anti or pro-gay tendency that should, in theory, correlate well with MHS scores.

3.2.2 Method

3.2.2.1 Participants

Participants for this experiment were recruited from either the NUI, Maynooth undergraduate population or from acquaintances of the experimenter via word of mouth recruitment. All participants were adults (over 18 years of age) and did not suffer from any condition that would interfere with their ability to perform routine cognitive tasks. As was the case with previous experiments, the current study imposed an exclusion criterion for those who failed to achieve fluency (i.e., ten trials in a row correct) within 200 trials on the first baseline block. Consequently, one participant was excluded from the current analysis. The remaining sample (n = 30) was comprised of 13 males and 17 females aged between 18 and 36 (M = 24.3; SD = 4.05). Additionally, it should be noted that all participants included in this analysis were either attending University or were recent University graduates.

3.2.2.2 Ethical Considerations

As discussed previously, all elements of this research process received prior approval from the NUI, Maynooth Ethical Committee, and were all in accordance with the Psychological Society of Ireland (PSI) code of ethics, and with the NUI, Maynooth Department of Psychology guidelines for safe working practice in psychological research.

3.2.2.3 Materials

3.2.2.3.1 The FAST Software Package and Stimulus Set

The version of the FAST utilised in the current experiment was identical to previous experiments, and was presented on a 13” laptop (1280x800 pixel resolution). Image stimuli for this experiment were again predominantly obtained from the IAPS database (Lang, et al., 1999). This is with the exception of the images representing the category ‘Gay’, which were purchased from an online stock photo database (www.istockphoto.com), in addition to a small selection of stimuli used during the practice block. Stimuli obtained from the IAPS

were matched roughly for standardised arousal ratings and valence across all blocks except for the consistent and inconsistent blocks during the testing portions of the FAST, in which one of the stimulus categories consisted of either positively or negatively valenced and highly arousing images (as rated by the IAPS manual). Images were cropped slightly to achieve a perfectly square shape and in a small number of cases to eliminate peripheral objects in the image view. All images were 150x150 pixels in size (at a resolution of 72 dpi). During all blocks, stimulus categories were comprised of four image exemplars. The alphanumeric category code, category description, and their corresponding IAPS/istockphoto image identifier numbers are listed below (Table 3.2.1).

Table 3.2.1: Pictorial exemplars employed in Experiment 3b

Alphanumeric	**Stimulus Category**	**IAPS Image Number/Image Description/istockphoto Image Number**
Phase 1: Practice block		
P1	Circles	Four discriminable images of circles in varying sizes and colours
P2	Squares	Four discriminable images of squares in varying sizes and colours
Q1	Chinese Characters	Four discriminable Chinese symbols
Q2	Numbers	Numbers 1, 2, 3 and 4 in black 48-type font
Phase 2: Baseline 1		
X1	Mushrooms	5531, 5532, 5533, 5544
X2	Cakes	7200, 7250, 7270, 7282
Y1	Birds	1333, 1419, 1450, 1740
Y2	Furniture	7025, 7235, 7705, 7710
Phase 3: FAST [1]		
A1	Gay men	1917875, 2007866, 3657240, 10197611
B1	Small animals	1440, 1460, 1710, 1750

N1	Nature	5201, 5220, 5760, 5780
N2	Vehicles	7130, 7140, 7595, 8510
Phase 4: FAST [2]		
A1	Disgust images	9300, 7380, 3261, 9373
B1	Gay men	1917875, 2007866, 3657240, 10197611
N1	Nature	5201, 5220, 5760, 5780
N2	Vehicles	7130, 7140, 7595, 8510
Phase 5: Baseline 2		
X3	Sky	5593, 5594, 5990, 5991
X4	Babies	2040, 2050, 2250, 2260
Y3	Food	7281, 7291, 7350, 7480
Y4	Zoo Animals	1601, 1661, 1810, 1812

3.2.2.5.3 The Modern Homonegativity Scale

The current experiment also utilised an identical 10-item MHS scale for the purposes of identifying levels of self-reported homonegativity among participants.

3.2.2.4 Design

The current experiment employed a similar correlational design as in Experiment 3a, whereby MHS scores will be correlated against FAST performance.

3.2.2.5 General Experimental Sequence

3.2.2.5.1 Informed Consent

The procedure for informed consent was identical to previous experiments.

3.2.2.5.2 Function Acquisition Speed Test

Upon supplying written consent, the participant was then required to complete the FAST portion of the experiment. This portion was identical to Experiment 3a, with the exception of the second test block, within which the two test stimulus categories (men

shaking hands and positive images) were replaced with two new categories (Gay men and images evoking disgust, such as vomit or a dirty toilet). The reinforcement contingencies were identical to Experiment 3a.

3.2.2.5.3 The Modern Homonegativity Scale

The MHS was again administered following the FAST portion of the experiment as a means of establishing participants' levels of explicit homonegativity, against which FAST scores could be compared.

3.2.2.5.4 Debriefing

Participants were fully debriefed regarding the nature of the experiment immediately upon completion in an identical manner to previous experiments.

3.2.3 Results

3.2.3.1 Modern Homonegativity Scale Scores

Table 3.2.2 below outlines the MHS scores for the current sample (M = 20.60; SD = 7.67). Given that possible scores on the MHS range from 10 to 50, with 50 being the highest, the current sample overall seems to have low-to-mid levels of explicit homonegativity, though the sample is slightly clustered around the higher end of the scale. Again, similar to Experiment 3a, this is not necessarily unexpected given explicit pro-gay bias that may be more prevalent in an educated University sample relative to the wider population (e.g., Cerny & Polyson, 1984; Elchardus & Spruyt, 2009; Ohlander et al., 2005).

Table 3.2.2: MHS scores for all participants

Participant	MHS	**Participant**	MHS
1	13	**16**	20
2	23	**17**	31
3	28	**18**	17
4	30	**19**	19
5	15	**20**	16
6	17	**21**	28
7	10	**22**	32
8	11	**23**	22
9	19	**24**	31

10	11	**25**	32
11	31	**26**	14
12	10	**27**	15
13	14	**28**	26
14	33	**29**	16
15	20	**30**	14

3.2.3.2 Practice Blocks

As addressed previously, the practice block data will not be discussed here as Phase 1 served only to familiarise participants with the general procedure.

3.2.3.3 Baseline Blocks

Table 3.2.3 documents the number of trials required by each participant to reach the fluency criterion across the two baseline blocks. An inspection of the data set reveals a similar pattern to that observed in previous experiments, with acquisition rates typically decreasing from the first to the second baseline block. A Wilcoxin Signed Ranks test revealed this to be a statistically significant reduction (z = -3.40, p = .002, two-tailed) with large effect size, again suggesting that participants are still improving substantially in terms of baseline acquisition rates as they progress through the FAST procedure.

Table 3.2.3: Number of trials to criterion across both baseline blocks (Phases 2 and 5. Note: the difference is obtained by subtracting the number of trials to criterion for the consistent block from the inconsistent block.

P	Baseline 1	Baseline 2	Mean Baseline	P	Baseline 1	Baseline 2	Mean Baseline
1	18	26	22	**16**	21	18	19.5
2	21	28	24.5	**17**	21	16	18.5
3	26	11	18.5	**18**	14	12	13
4	25	17	21	**19**	22	11	16.5
5	14	26	20	**20**	31	27	29
6	10	12	11	**21**	16	18	17
7	12	10	11	**22**	29	11	20
8	25	16	20.5	**23**	33	20	26.5
9	26	26	26	**24**	28	14	21
0	19	14	16.5	**25**	50	53	52
11	20	11	15.5	**26**	50	11	30.5
12	24	10	17	**27**	44	26	35
13	44	26	35	**28**	22	26	24
14	41	17	29	**29**	11	17	14
15	77	13	45	**30**	36	32	34

3.2.3.4 Function Acquisition Speed Test

3.2.3.4.1 Acquisition Rates

Table 3.2.4 documents the acquisition rates for the inconsistent and consistent blocks for the Homopositivity FAST. Wilcoxin Signed Ranks tests comparing the inconsistent and consistent blocks revealed there to be no significant difference in acquisition rates between the two blocks for the Homopositivity FAST (z = -1.420, p = .156), suggesting no FAST effect at a group level.

Table 3.2.4: Number of trials to criterion across both FAST blocks for the Gay-Good block (Phase 3/4).

P	Inconsistent	Consistent	Difference	P	Inconsistent	Consistent	Difference
1	35	14	21	**16**	16	15	1
2	15	14	1	**17**	21	20	1
3	23	20	3	**18**	14	38	-24
4	11	15	-4	**19**	35	15	20
5	13	10	3	**20**	26	27	1
6	17	13	4	**21**	18	12	6
7	14	12	2	**22**	11	20	-9
8	33	23	10	**23**	28	16	12
9	18	20	-2	**24**	15	109	-94
10	39	22	17	**25**	30	57	-17
11	23	19	4	**26**	25	11	14
12	13	13	0	**27**	40	20	20
13	22	19	3	**28**	10	19	-9
14	18	33	-15	**29**	19	17	2
15	30	49	-19	**30**	17	35	-18

Table 3.2.5 below documents the number of trials to criterion for the Homonegativity FAST. Wilcoxin Signed Ranks tests revealed there to be no significant difference in acquisition rates between the consistent and inconsistent blocks for the Homonegativity FAST (z = -1.420, p = .156), suggesting no FAST effect at a group level.

Table 3.2.5: Number of trials to criterion across both FAST blocks of the Gay-Disgust block (Phase 3/4).

P	Inconsistent	Consistent	Difference	P	Inconsistent	Consistent	Difference
1	24	37	-13	**16**	13	12	1
2	19	13	6	**17**	28	14	14
3	34	55	-21	**18**	36	16	20
4	37	25	12	**19**	28	24	4

5	11	13	-2	**20**	13	21	-8
6	35	16	19	**21**	45	13	32
7	16	14	2	**22**	10	12	-2
8	19	14	5	**23**	63	10	53
9	24	11	13	**24**	15	20	-5
10	20	26	-6	**25**	40	50	-10
11	40	15	25	**26**	10	12	-2
12	22	14	8	**27**	10	22	-12
13	20	19	1	**28**	45	31	14
14	51	20	31	**29**	16	28	-12
15	22	77	-55	**30**	40	10	30

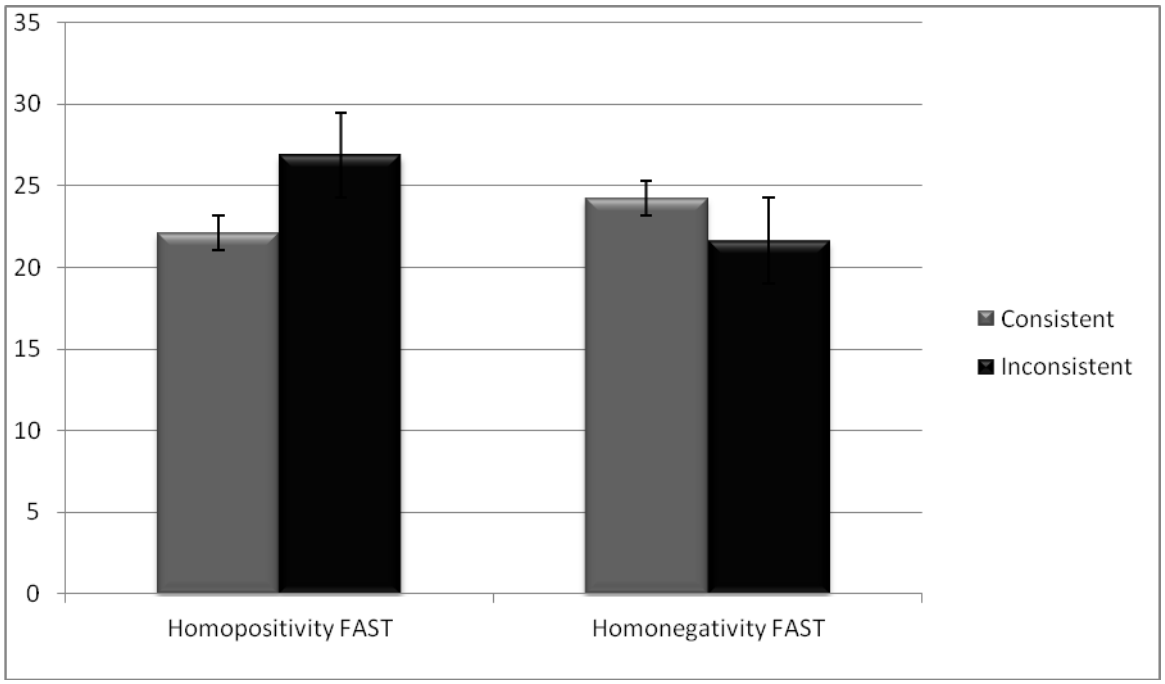

Figure 3.2.1: Mean trial acquisition rates for the consistent and inconsistent blocks for the Homopositivity and Homonegativity FASTs

3.2.3.4.2 Strength of Relation Indices

The current experiment calculated a simple SoR index for each participant for both the Homopositivity (Table 3.2.6) and Homonegativity FASTs (Table 3.2.7), again according to the formula outlined by O'Reilly and colleagues (O'Reilly et al., 2012). An examination of the SoR indices below reveals a varied sample in terms of both Homopositivity SoRs (*M* = -1.73, *SD* = 15.01), which are highly skewed toward lower scores and has high numbers of extremes, and Homonegativity SoRs (*M* = 4.74, *SD* = 15.01), which are slightly skewed

toward the lower end though there are few extreme cases. No particularly strong trends emerge upon a first inspection of the data. There does appear to be a tendency, however, toward weaker and more negative SoRs for the Homopositivity FAST scores and stronger, more positive SoRs for the Homonegativity FAST scores, although a Wilcoxin Signed Ranks test investigating the difference between both sets of scores revealed this difference to be non-significant (z = -1.568, p = .117). Further to this, a single-sample t-test comparing both sets of scores to zero revealed no significant overall effect for either the Homopositivity, t (29) = 1.601, p = .554, or the Homonegativity SoRs, t (29) = 1.730, p = .094, further suggesting that the sample is not particularly biased in either a homopositive or homonegative direction.

Table 3.2.6: Strength of Relation (SoR) Indices for each participant for the Gay-Good FAST (Phase 3/4). Note: large positive SoRs indicate a strong pre-existing relation, while zero or negative SoRs suggest a weak or absent pre-existing relation

Participant	[SoR]	**Participant**	[SoR]
1	15.7	**16**	0.81
2	0.72	**17**	0.79
3	2.36	**18**	-21.62
4	-3.03	**19**	16.39
5	2.3	**20**	-.70
6	3.85	**21**	4.90
7	3.85	**22**	-6.92
8	7.63	**23**	8.45
9	-1.42	**24**	-71.12
0	11.18	**25**	-15.78
11	2.6	**26**	9.45
12	0	**27**	12.90
13	1.95	**28**	-6.52
14	-8.9	**29**	1.74
15	-11.52	**30**	-11.90

Table 3.2.7: Strength of Relation (SoR) Indices for each participant for the Gay-Disgust FAST (Phase 3/4).

Participant	[SoR]	**Participant**	[SoR]
1	-9.7	**16**	0.81
2	4.32	**17**	11.02
3	16.50	**18**	18.01
4	9.10	**19**	3.28
5	-1.54	**20**	-5.48

6	18.27	**21**	26.01
7	1.92	**22**	-1.53
8	3.82	**23**	-3.79
9	14.90	**24**	-5.85
0	-3.95	**25**	-1.35
11	16.23	**26**	-5.40
12	6.50	**27**	-7.90
13	.65	**28**	10.14
14	21.23	**29**	-10.44
15	-46.06	**30**	19.87

While each of the two FAST blocks were designed to provide an absolute strength of relation index between 'Gay' and 'Good' and between 'Gay' and 'Disgust', it was also decided that it may be useful to calculate a combined SoR index that assessed participants' overall levels of implicit homonegativity, as measured by the two FAST exposures. Thus, a simple combined Homonegativity SoR index (Table 3.2.8) was calculated for each participant by subtracting the Homopositivity SoR from the Homonegativity SoR, and dividing this by two. The distribution of combined SoR indices is highly varied among the sample as a whole (M = 4.31; SD = 14.76) with a high number of extreme cases, though this is to be expected given the variance in self-reported levels of homonegativity. The data is skewed quite strongly toward positive values, which would suggest an overall trend in the data toward homonegative bias, however, a single-sample t-test was conducted to investigate whether the combined SoR scores were significantly different to zero revealed no significant group effect; t (29) = 1.601, p = .120. However, a data inspection does reveal a trend in combined SoR values, with those with particularly high MHS scores (around 30 or higher) appearing to have stronger, more positive SoRs, indicative of homonegativity, while those with the lowest MHS scores (around 15 or lower) appear to have weaker and more negative SoRs, indicative of strong homopositivity.

Table 3.2.8: Combined Strength of Relation (SoR) Indices for each participant (Phase 3/4).

Participant	[SoR]	**Participant**	[SoR]
1	-12.35	**16**	0
2	2.52	**17**	5.12
3	7.07	**18**	19.82
4	6.07	**19**	-6.56
5	-1.92	**20**	-4.80
6	7.21	**21**	10.56
7	-0.97	**22**	2.70
8	-1.91	**23**	14.43
9	8.16	**24**	67.33
10	-7.57	**25**	4.97
11	6.82	**26**	-5.40
12	3.25	**27**	-7.90
13	-0.65	**28**	8.33
14	15.07	**29**	-6.09
15	-17.27	**30**	15.88

3.2.3.4.3 Correlational analyses

It was expected that the MHS would correlate generally with SoRs on both FASTs. To investigate this, a simple analysis was conducted examining the strength of the correlation between MHS scores for the entire sample and with the Homopositivity, Homonegativity and Combined SoRs indices. Spearman's Rho correlational analyses revealed a strong, positive correlation with Combined SoRs (r = .533, p = .002), a medium, negative correlation with the Homopositivity SoRs (r = -.489, p = .006), and a non-significant positive correlation with the Homonegativity SoRs (r = .306, p = 1.0). This can be seen graphically in the scatter plots below, which reveal a strong negative linear relationship between MHS and Homopositivity SoRs (Figure 3.2.1), a moderately positive linear relationship between the MHS and Homonegativity SoRs (Figure 3.2.2), and a strong positive linear relationship between the MHS and the Combined scores (Figure 3.2.3). Thus, in effect, correlations were in the predicted direction.

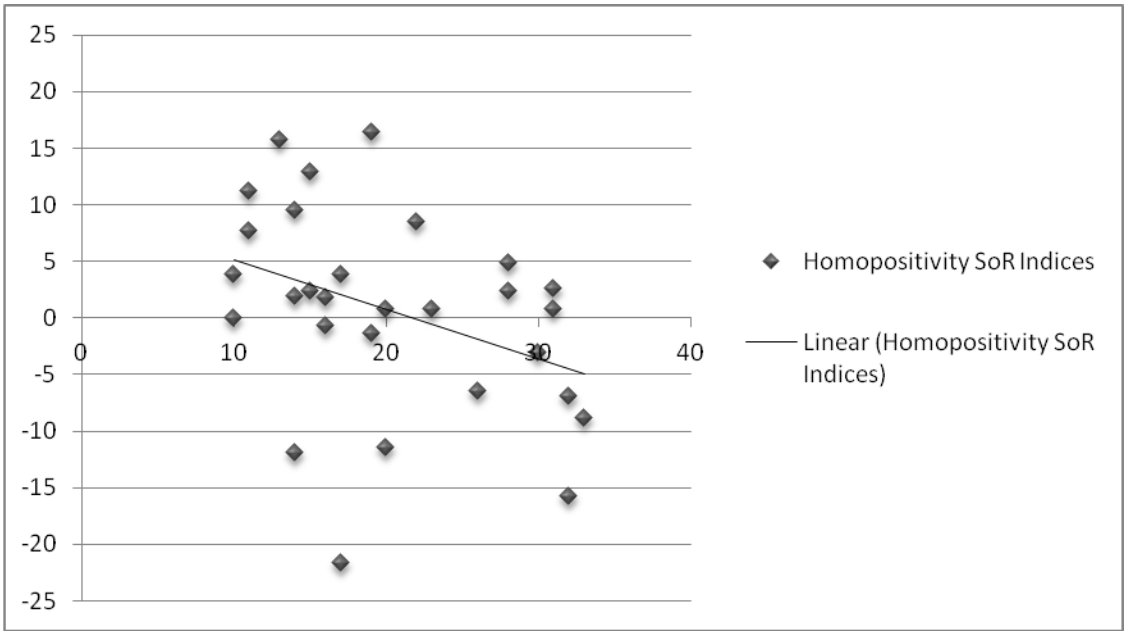

Figure 3.2.1: Correlations between the MHS and Homopositivity SoR Indices

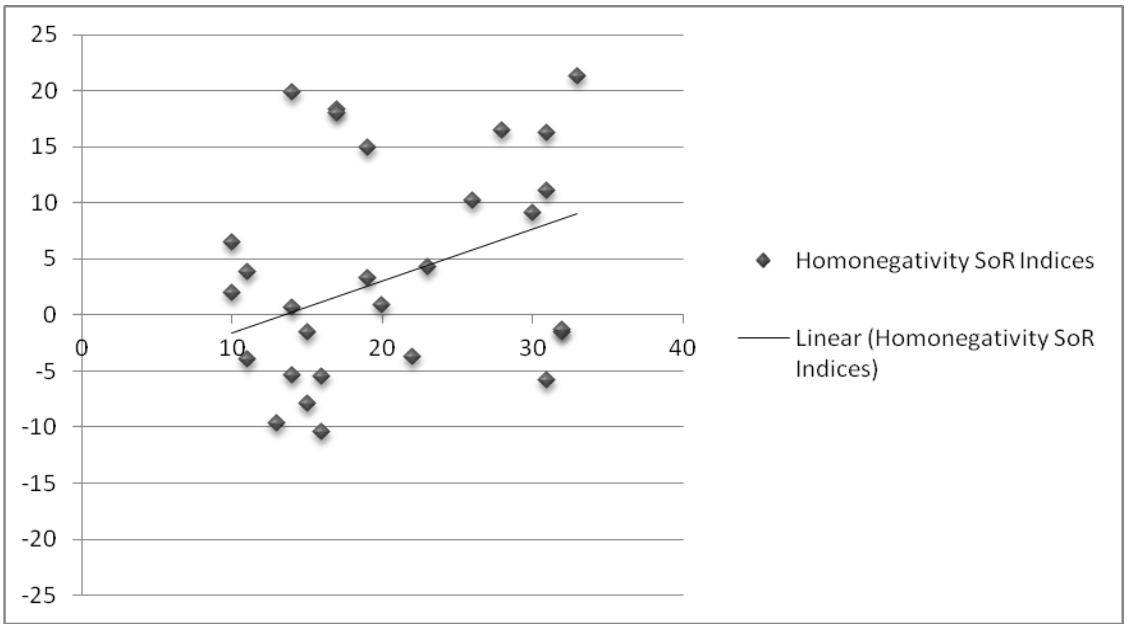

Figure 3.2.2: Correlations between the MHS and Homonegativity SoR Indices

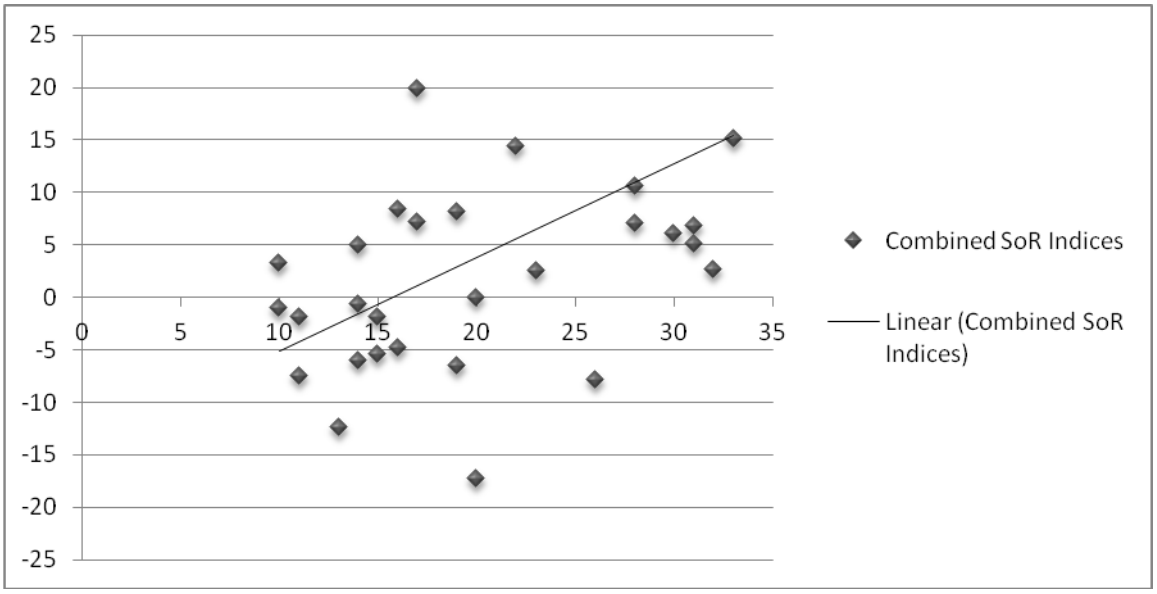

Figure 3.2.3: Correlations between the MHS and Combined SoR Indices

It was also necessary to examine the strength of the correlation between the two FASTs as well. As mentioned in the introduction to this experiment, the inclusion of a second Homonegativity FAST was to serve as a 'qualifier' of sorts alongside the Homopositivity FAST. That is, the Homopositivity FAST score could be considered as more meaningful when it's score is contrasted with a Homonegativity FAST examining the opposite stimulus relation. In this way, it was not expected that the two scores should necessarily mirror each other (i.e., that there would be a strong, negative correlation between them), however, it was hypothesised that scores should not *positively* correlate. To test this, a simple Spearman's Rho correlation was conducted between the two sets of scores, which revealed a weak, negative correlation (r = -.063, p =.71) that was non-significant.

3.2.3.4.4 Order Effects

It was also important to investigate any potential order effects in the data set, that is, whether the outcomes on the FAST were influenced by the order in which the consistent and inconsistent blocks were administered. To do this, a Mann-Whitney U test was conducted comparing trial differentials for the inconsistent and consistent blocks within both FASTs.

These analyses revealed no significant differences in the current data based on order, for either the Homopositivity (z = -1.379, p = .168, two-tailed) or Homonegativity FAST (z = -1.165, p = .244, two-tailed).

Table 3.2.10: Running order for inconsistent and consistent test blocks (1: Inconsistent first, 2: Consistent first) for the Gay-Good block

Participant	Order	**Participant**	Order
1	2	**16**	2
2	2	**17**	2
3	1	**18**	1
4	1	**19**	1
5	1	**20**	1
6	2	**21**	2
7	2	**22**	1
8	1	**23**	2
9	1	**24**	2
0	1	**25**	1
11	2	**26**	1
12	2	**27**	2
13	1	**28**	1
14	1	**29**	1
15	1	**30**	1

Table 3.2.11: Running order for inconsistent and consistent test blocks (1: Consistent first, 2: Inconsistent first) for the Gay-Disgust block

Participant	Order	**Participant**	Order
1	1	**16**	2
2	2	**17**	2
3	1	**18**	2
4	1	**19**	1
5	2	**20**	2
6	1	**21**	2
7	2	**22**	1
8	1	**23**	1
9	2	**24**	1
0	2	**25**	1
11	2	**26**	1
12	1	**27**	2
13	2	**28**	1
14	2	**29**	1
15	1	**30**	2

Table 3.2.12 below documents the order in which the two FASTs were administered for each participant. To investigate any order effects based on this order, a Wilcoxin Signed

Ranks test comparing SoRs between participants' first and second FAST revealed no statistically significant difference between SoRs for FAST blocks [1] and [2] (z = -1.566, p = .117, two-tailed), suggesting no major order effects in the current data.

Table 3.2.12: Running order for FAST blocks (1: Gay-Disgust first, 2: Gay-Good first)

Participant	Order	**Participant**	Order
1	1	**16**	2
2	2	**17**	1
3	2	**18**	2
4	1	**19**	1
5	1	**20**	2
6	1	**21**	2
7	2	**22**	2
8	2	**23**	1
9	2	**24**	2
0	1	**25**	2
11	1	**26**	1
12	1	**27**	2
13	2	**28**	2
14	2	**29**	2
15	1	**30**	2

3.2.3.6 Summary

This experiment again revealed a significant reduction in baseline acquisition rates across the two baseline blocks for all participants, providing further evidence for a problematic practice effect inherent in the FAST procedure. With regard to the two FAST testing blocks, the trials acquisition rates for the inconsistent and consistent blocks were not significantly different to one another for either the pro-gay or the anti-gay FAST, suggesting no FAST effect for the sample overall. Further to this, SoR distribution within the current sample reflected a similar trend, with no significant effect against zero found for the Homopositivity, Homonegativity, or Combined SoR indices. Analyses also found there to be no significant difference between the Homopositivity and Homonegativity SoR values, which would suggest no strong effect at a group level. Spearman's Rho correlational analyses, however, revealed that increases in MHS scores are associated with significantly stronger and

more positive Combined SoRs, with significantly weaker and more negative Homopositivity SoRs, and with non-significantly more negative Homonegativity SoRs. This would indicate, therefore, a significant relationship between the FAST and self-reported levels of homonegativity. Finally, this experiment found no order effects in the current data. That is, there was no significant difference in acquisition rates based on the order in which the inconsistent and consistent blocks were administered within both FASTs, as well as based o the overall order of the Homopositivity and Homonegativity FASTs.

3.2.4 Discussion

The aim of the current experiment was to investigate the effectiveness of a two-phase multi-exemplar pictorial FAST at detecting levels of Homonegative and Homopositive bias in a University-educated population. It was expected that there would be some general correlations between self-reported levels of homonegativity (i.e., MHS scores) and FAST scores, such that increases in MHS scores would correlate with increases in homonegative bias as determined by the FAST. More specifically, it was expected that increases in MHS scores would be generally correlated with stronger, more positive SoRs (i.e., greater behavioural resistance) for the Homopositivity FAST and with relatively weaker, more negative SoRs for the Homonegativity FAST (i.e., greater behavioural facilitation). Unlike Experiments 1 and 2, however, which employed more diverse samples, it was not expected that there would be a trend toward homonegative FAST scores for the sample as a whole.

Following statistical analyses, it would appear that the current experiment did show trends in expected directions. Firstly, there was no significant homonegative or homopositive bias for the sample as a whole, which was expected given the tendency toward low self-reported homonegativity, in addition to the sample being recruited from a University-educated population. However, there was a strong significant correlation between self-reported homonegativity (as ascertained by the MHS) and scores on the FAST, with a particularly strong correlation between the MHS and the Combined SoR indices. Indeed, the correlation coefficient obtained here between the MHS and the Combined score (r = .533) is similar to the coefficients obtained in previous IAT studies of Homonegativity (e.g., .62; Banse et al., 2001). Thus, it may be cautiously inferred that the stimuli used here were appropriate at depicting the target constructs for a functionally understood two-phase Homonegativity/Homopositivity FAST. That is, they may have been useful in detecting participants' verbal histories of stimulus relations. This would lend support, therefore, to the

idea that affective pictorial exemplars are more suitable for an affective construct such as anti or pro-gay biases than more semantically representative verbal or pictorial stimuli.

While previous research has suggested that these exemplars would be more appropriate than those used in Experiments 1 and 2, thereby providing support for the above assumption, there are some alternative interpretations of the overall findings that should also be considered. Firstly, it may be possible that the increased salience of the target categories (i.e., 'Gay', 'Good', and 'Disgusting') relative to the neutral categories in the baseline and target blocks unintentionally influenced results. Previous IAT research has suggested that the IAT effect may be mediated by the salience of the target stimuli relative to the distracter stimuli (e.g., Rothermund & Wentura, 2004), which are typically relatively neutral in evaluative functions. That is, the two salient target stimuli may just be associated (i.e., they may form a natural response class) for an individual because they are both salient relative to the two neutral categories, rather than because they were previously related. This could have been a potential confound here, given that the target categories in the FAST blocks were considerably higher in affective valence than the neutral stimuli, and thus it is possible that they may have naturally formed a response class based on salience alone. However, no effects were observed here that suggest that the consistent blocks (i.e., the blocks within each FAST that require the formation of an A-B response class) were inherently easier than the inconsistent blocks, which would be expected if the salience of the categories influenced performance. In fact, such an outcome would be easily spotted and would in fact render at least one of the FAST performances completely unrelated to MHS scores. More specifically, given a particular verbal history of relating Gay stimuli to positive or negative stimuli, a participant in Experiment 3b should require a different number of training trials in order to complete the consistent blocks on both FASTs. That is, one consistent block requires them to form a functional response class for Gay and Positive stimuli, whereas the other requires

them to learn a functional response class for Gay and negative stimuli. Put simply, the participant's history is unlikely to facilitate or retard learning on both blocks. In fact, the primary advantage of a two FAST test is that it provides the researcher with two opportunities to measure the same bias in different 'directions' (i.e., Gay-positive and Gay-negative). The outcomes of both tests should be different for each participant, and the magnitude of the combined score (which measures an overall bias direction) should correlate with other measures of bias (e.g., the MHS), which it did in this experiment. By definition, therefore, the two consistent blocks on the FAST tests did not *both* produce lower than expected trial requirements (i.e., were easy), because if this had been the case no biases would have been found at all (i.e., the two FAST outcomes would have contradicted and cancelled each other out in the calculation of the combined SoR score). Thus, it would appear that the salience of the images is not a significant issue, although this would of course require substantiating in further research.

One additional potential point of concern may be that any enhanced familiarity with the target stimuli may contribute to the learning bias observed in the FAST. Previous IAT research revealed that participants' enhanced familiarity with target stimuli may influence IAT scores, such that a higher degree of familiarity with certain stimuli (e.g., White faces) may lead to shorter response latencies for those over less familiar stimuli (e.g., Black faces; Rothermund et al., 2005). Thus, it is possible that a similar issue affected outcomes here. Initial research concerned with the FAST's procedural development focused almost entirely on the use of nonsense syllables, whereby matching-to-sample training would train a particular A-B relation, the strength of which would then be assessed using the FAST methodology employed here. However, it may be that the increased familiarity with the A and B stimuli relative to the neutral N stimuli, engendered by extensive matching-to-sample training using the A and B nonsense words, may in fact lead to the target stimuli forming a

natural response class (i.e., ‘things I am familiar with’). This is a potential confound that research will have to explore. The general point, however, may also apply in this case, even though no classes at all were trained in the laboratory prior to the FAST. Thus, the possibility that FAST effects are due to some unidentified but pre-existing familiarity with target stimuli relative to the N stimuli should of course be duly considered. However, when designing the current experiment, all of the stimulus categories were selected randomly and were roughly matched for perceived familiarity. That is, each of the stimulus categories were selected from frequently encountered everyday constructs (i.e., ‘Vehicles’, ‘Nature’) that should not have been less familiar to participants than the target constructs. While it is always a possibility that specific stimulus categories may form a pre-existing natural class for individual participants, it is unrealistic to attempt to control for all possible extraneous stimulus functions based on the idiosyncratic learning histories of individual participants once one has committed to building an ecologically valid test for real world applications using naturalistic stimuli. Of course, there were no significant trends in the data set that would indicate a familiarity effect (i.e., whereby the consistent block would be consistently easier than the inconsistent for all participants) for either FAST. As per the issue of salience, any such effect would have quickly undermined a two-FAST test by creating contradictory FAST outcomes in each FAST and therefore leading to a Combined SoR score tending towards zero and not correlating with other dependent measures (e.g., the MHS). This did not happen in the current study, and so any effect based on pre-existing idiosyncratic cross-class relations was relatively small, albeit still potentially real. Additionally, it should be noted at this point that the use of multiple exemplars also reduces the potential impact on data of any pre-existing cross-class relations, insofar as it reduces the number of trials on which such confounded learning trials are presented and, therefore, the overall impact of those confounds on overarching learning rate of participants across the two FAST blocks as a whole.

It would appear that the most coherent and parsimonious interpretation of the current effects would be that the current exemplars were effective in generating two differing rates of learning across each of the two pairs of blocks in the Function Acquisition Speed Tests. Working from this assumption, it is interesting to note more nuanced data effects within the two-phase FAST when the findings are examined in more detail. Specifically, MHS scores correlated strongly and significantly with SoR indices for the Homopositivity FAST, however they did not correlate significantly with the SoRs for the Homonegativity FAST. In other words, as MHS scores increased, Homopositivity SoRs significantly decreased, suggesting a strong correlational relationship. However, increases in Homonegativity SoR indices were not significantly correlated with increases in MHS scores, though it was in the expected positive direction. Additionally, correlational analyses revealed no significant relationship between these two sets of scores, such that having a high score on one of the FASTs was not associated with a low score on the other. Thus, this experiment revealed that it does not necessarily follow that if an individual is strongly pro-gay according to one FAST that they will be very weakly anti-gay on the other. On the other hand, it is a validation of the FAST method that there was not a positive correlation between these two measures, indicating that scores were at least not in the same direction on both FASTs. Nevertheless, the important finding here is that not only are the two FASTs not bidirectional in terms of effects (i.e., their scores do not perfectly mirror each other in directionality), but that the Homopositivity FAST seems to be contributing most significantly to the overall combined score effect.

There are many possible interpretations to this finding, the first of which being that, conceptually, it may be the case that the strength of an individual's pro-gay verbal history is not necessarily the exact inverse of their anti-gay verbal history. In other words, it may be possible that being strongly anti-gay does not correlate with being strongly NOT pro-gay (i.e., the inverse of pro-gay), and vice versa. However, as discussed in Chapter 1, previous

IAT and IRAP research has predominantly focused on identifying preferential pro/anti attitudes towards a particular construct in relation to another, such as, for example, pro-slim/anti-fat (e.g., Roddy, Stewart, & Barnes-Holmes, 2010), pro-Black/anti-White (e.g., Barnes-Holmes et al., 2010; Heider & Skowronski, 2007), or pro-gay/anti-Straight (e.g., Jellison et al., 2004) preferences. That is, while it may be possible that the current findings suggest a disparity between positive and negative attitudes toward a particular construct, the current study's approach looking at a single construct individually from a pro and anti perspective is a relatively novel methodology. Indeed, a review of the literature suggests that this study is the first of its kind to directly examine a single construct from both a pro and anti perspective. Thus, while it may be possible that the current findings are due to attitudes are not bi-directional (i.e., having a strong pre-existing Gay-Good verbal history does not correlate with a reversed weak Gay-Bad history exactly), this would need to be substantiated in further research within a similar pro/anti paradigm.

An additional interpretation to this finding could be that it resulted from the current FAST software employed in this experiment. As described in the introduction to this thesis, the FAST is conceived as a measure of behavioural resistance to the formation of functional stimulus classes (e.g., Gay-Good) that may be orthogonal to an individual's verbal histories. As such, this thesis has predominantly focused on assessing the degree of resistance to the formation of a functional Gay-Good stimulus class rather than the degree of facilitation to the formation of a Gay-Bad class. Thus, as discussed above, while it is possible that the complexities of human verbal behaviour may have resulted in the contrasting verbal histories not being bi-directional (i.e., a strong pre-existing Gay-Good stimulus class is not the exact inverse of a weak pre-existing Gay-Disgusting stimulus class), another interpretation could be that it relates to a methodological issue whereby the FAST can more clearly measure resistance relative to facilitation. That is, it may be easier to assess the extent to which

stimulus class formation for two sets of target categories is retarded relative to a baseline (i.e., two neutral blocks) than it is to for assess how that behaviour is facilitated. This is simply because the task is at heart relatively easy to complete, and baseline trial requirements are generally low (in the current experiment, average baseline acquisition rates were typically between 20-40 trials). Thus, facilitated learning is by definition limited in magnitude, insofar as any increase in learning rates over the baseline learning rate will quickly be obscured by a "floor effect". In contrast, a retarded learning rate can lead to a trial requirement many times greater than the baseline learning rate and as a result, the FAST metric is intrinsically more sensitive to learning delay over learning facilitation. In other words, if the baseline is already quite low, behavioural facilitation is not as evident as resistance in learning. Thus, it does not necessarily mean that in cases where little resistance is encountered that the pre-existing stimulus relation is significantly weaker *per se*, but rather that it is detected less easily than resistance by the current FAST design.

The current research provided empirical support for this interpretation. As expected, the Homopositivity FAST correlated significantly with MHS scores, suggesting that the greater an individual's explicit homonegativity, the greater resistance will be encountered when forming a Gay-Good stimulus class. In contrast, as scores increased on the MHS, forming a functional stimulus class for the categories 'Gay' and 'Disgusting' was facilitated, though not at a significant level relative to the baseline. In other words, the data revealed that the behavioural resistance observed in the Homopositivity FAST as scores on the MHS increased was significantly different to baseline acquisition rates, while these increases were not correlated with significant behavioural facilitation for the Homonegativity FAST. Thus, the level of behavioural resistance was more pronounced in terms of FAST scores than the corresponding behavioural facilitation.

This does raise certain issues, however, regarding the lack of significant resistance observed to the formation of a functional Gay-Disgusting response class as scores on the MHS decreased. While there was a positive correlation observed between MHS and Homonegativity FAST scores, in theory, it would have been expected that similar levels of resistance would have encountered by the strongly pro-gay individuals for the Homonegativity FAST as was encountered by the strongly anti-gay individuals on the Homonegativity FAST. That is, following the same logic as above, there should have been a significant positive correlation between on the MHS scores and the Homonegativity FAST. However, as discussed previously in Chapter 2, given that the MHS is a self-report measure, it was suggested the MHS scores may be less sensitive indicators of levels of low anti-gay bias and especially pro-gay bias and, consequently, more variability in the relationship between low MHS scores and FAST scores may be expected. Thus, one possible interpretation could be that there was greater variance in homonegativity at the lower end of the scale, which may have accounted for the weaker behavioural resistance observed to the formation of a Gay-Disgusting stimulus class.

While the issues above undoubtedly require further research and consideration, it should be noted that the primary rationale for the inclusion of a second Homonegativity FAST in this experiment was to qualify the findings of the Homopositivity FAST. That is, it served to provide some support to the findings from the Homopositivity FAST by showing non-identical or even contrasting outcomes. In other words, the two-phase FAST should be a more stable representation of an individual's verbal history, given that a strong Gay-bad relation should not in theory correlate with a strong co-existing Gay-Good relation. In this way, the Combined SoR index that was calculated should have been an even more stable and reliable score than the two individual FASTs taken independently. Indeed, the current findings support this assumption, with the strongest correlation observed emerging between

the MHS and Combined SoR indices. Thus, even though each individual FAST score provides an independent and absolute measure of an individual's Gay-Good or Gay-Disgusting verbal histories, it would appear that, when combined, they each serve to further inform the other.

It is also important to note here that the use of a combined score is not unlike the IAT and IRAP strategies. However, in both cases, the overall scores obtained within these measures are necessarily relative. Both measures typically provide a combined score that averages, for example, an individual's pro-gay score with their pro-straight score to provide a more stable representation (see Greenwald et al., 2003, and Barnes-Holmes et al., 2010, for more detailed reviews of the IAT and IRAP scoring strategies, respectively). That is, within both measures, the aim is typically to provide a level of preference of one construct over another, such as between Black and White individuals, rather than to identify an absolute bias in stimulus class formation (i.e., a learning bias). Thus, while a two-FAST approach appears to mimic a relativistic approach to attitude measurement the important key difference is that each FAST within the pair provides an absolute measure of the bias of interest in its own right. The Combined SoR, therefore, remains an absolute measure, given that it reflects a double measure of the bias in two different directions, so that the results of each confirm the findings of the other. However, it should be noted at this point that previous unpublished social-cognitive research has attempted to develop a non-relative alternative to the original IAT, known as the Single Target IAT (ST-IAT; Wigboldus, Holland, & van Knippenberg, 2004) . Essentially, this measure addresses the relativism inherent in the IAT by presenting a single target in isolation alongside two positive and negative attributes. However, it should also be taken into account that the ST-IAT still relies on the similar assumptions, procedure, and scoring format as the standard IAT, regarding which there have been several significant criticisms (e.g., Gavin et al., 2008; O'Reilly et al., 2012). Thus, the FAST achieves the same

outcome as was intended by the ST-IAT, though from within a functionally understood paradigm.

Another important finding that should be discussed here was the absence of any particular homopositive or homonegative trend overall in the data. Although it was expected that there would be no strong homonegative trend given the MHS scores and that the sample was young and University-educated, it is important to consider the possibility that society, while not necessarily moving away from anti-gay attitudes on the whole, may in fact be progressing away from more traditional forms of homonegativity (i.e., those based on more direct objections to homosexuality). Previous research (e.g., Morrison & Morrison, 2011) noted such a shift in terms of self-reported levels of homonegativity from traditional to more modern forms. However, this current experiment may indicate that this shift is occurring at a private level as well, at least within a young educated Irish population. That is, there may be a substantial shift occurring with regard to societal beliefs toward Gay individuals, and not just in the beliefs that people are typically willing to disclose. It may be possible, therefore, that a disgust-based approach may not be targeting the more modern and subtle forms of anti-gay bias among an Irish population. Indeed, another possible interpretation of the weaker correlation between MHS scores and the Homonegativity FAST relative to the Homopositivity FAST could be that the disgust images were not as representative as they could have been of modern homonegative attitudes in Ireland and, as such, failed to engender significant behavioural resistance to response class formation among strongly pro-gay individuals. Thus, these findings raise further questions regarding the nature of homonegative prejudice in Ireland that can be more thoroughly addressed in future research, and which will be revisited in the next Chapter.

Lastly, it should also be noted here that the current experiment again observed significant practice effects across the FAST blocks. As in all previous experiments reported

here, there was a significant reduction in the number of trials required to achieve fluency from the first to the second baseline blocks, suggesting that participants' performances were significantly improving as they progressed through the procedure. While no order effects were observed for the test blocks within either FAST, nor indeed between the Homopositivity and Homonegativity FASTs, practice effects may have affected data adversely. Thus, it is of considerable importance that this problem is sufficiently addressed in future research, and this matter will also be revisited in the final Chapter.

In conclusion, it would appear that the two-phase Homopositivity/Homonegativity FAST with highly affective pictorial stimuli is an effective "beta-version" assessment of homonegative attitudes. Scores from an explicit measure of anti-gay attitudes (the MHS) correlated significantly with scores on the FAST, indicating that this current FAST is a valid measure of homonegativity, at least within an Irish population. The wider implications of this research and the prospects for future experimentation will be discussed in detail in Chapter 4.

Chapter 4

General Discussion

4.1 Research Summary and Main Findings

This thesis set out to develop a functionally understood test methodology for anti-gay prejudice, and to provide an alternative to the standard explicit questionnaire or the eponymous Implicit Association Test (IAT; Greenwald et al., 1998). Specifically, this research aimed to develop the behaviour-analytic implicit measure known as the Function Acquisition Speed Test (FAST; O'Reilly et al., 2012), into a valid and reliable measure of homonegative attitudes. The FAST was developed based on the finding that a person's verbal histories of relating two classes of verbal stimuli may interfere with the formation of new relations that oppose or contradict the existing ones (Watt et al., 1991). That is, it built on the idea that the stronger a pre-existing relation is between two classes of stimuli (e.g., 'Gay' and 'Bad'), the more resistance will be encountered when learning to form a new class that is in opposition to it (e.g., 'Gay' not going with 'Bad'). Additionally, the FAST was developed as a means to measure not only relations that may have been formed directly in an individual's experiential history, but also perhaps those that may have been indirectly derived (i.e., through derived relational responding). This latter potential was not explored empirically in this thesis, but has been demonstrated recently by O'Reilly and colleagues (O'Reilly et al., in press). In this way, the FAST might be considered an assessment of "implicit" anti-gay bias, such that it may provide a means to assess an individual's verbal histories of relations between two classes of stimuli. As suggested in Chapter 1, in the event that a relation between two given classes has been derived, the FAST may be considered to measure what psychologists might formerly have referred to as "unconscious".

4.1.1 Exemplar Selection

The aim of the current research was to develop a Homopositivity (i.e., a pro-gay) FAST that would enable a precise empirical assessment of the degree of behavioural resistance encountered when learning to form a functional Gay-Good response class. As

such, the primary focus of the research has been to establish the optimal stimulus set for such a measure, given the noted importance of stimuli in modulating implicit test outcomes (e.g., Bluemke & Friese, 2005; De Houwer, 2002). Chapter 2 of this thesis, therefore, focused on determining the most appropriate stimulus type for a Homopositivity FAST by investigating the relative abilities of verbal and pictorial exemplars in representing the target constructs 'Gay' and 'Good'.

Firstly, Experiment 1 investigated the effectiveness of multi-exemplar verbal stimulus categories in a FAST for homopositive bias. Within this experiment, the degree of behavioural resistance to forming a functional response class between these two classes of stimuli was assessed using descriptive words representing the two constructs of interest (e.g., *Gay, Homosexual, Love, Happiness*) as exemplars. The effectiveness of these stimuli was assessed largely through the use of a 'known-groups' paradigm (e.g., Nettler & Golding, 1946), which involved distinguishing between participants on the basis of their self-reported homonegativity (ascertained by the explicit MHS scale) and investigating whether there were any significant differences in terms of behavioural resistance between these two groups in the expected direction (i.e., enhanced resistance observed for strongly anti-gay individuals). Within the IAT literature, word-IATs are typically associated with higher effect sizes than picture IATs (Foroni & Bel-Bahar, 2010), and consequently, it was expected that the word-FAST would engender differences in learning rates across test blocks for the two groups. However, this experiment yielded none of the predicted effects, with no particular trend observed in any direction.

Thus, for Experiment 1, it can be assumed that the verbal Homopositivity FAST was an ineffective measure of pro-gay attitudes. That is, it failed to identify the stimulus relation histories of these individuals. Based on previous research, however, it was conjectured that there might have been several reasons for this outcome, the most likely of which being that

verbal exemplars are simply not emotionally evocative for these particular constructs. Research suggests that anti-gay attitudes have a strong affective component (Fleming, 1967; Ostrom, 1969; Millar & Tesser, 1990). Simply put, this means that they may have emotional as well as other response functions transforming in accordance with the verbal network that constitutes the attitude. Therefore, it may be possible that semantic-verbal representation may not be as emotive as more perceptual pictorial exemplars in a FAST examining anti-gay prejudice. Thus, it was posited that while there may be little resistance to the formation of verbal classes that are incongruous with the attitude, there may be more behavioural resistance to the formation of functional stimulus classes involving emotive stimuli that are incongruous with the emotional stimulus functions that form part of the attitude.

An additional reason for the results of Experiment 1 may be related to cross-class overlap due to the exemplars' possible conflicting stimulus functions. It was discussed in Chapter 2 that exemplars should be selected that are unambiguous members of the stimulus class on interest, so that any potential cross-class overlap with other confounding constructs is minimized (De Houwer, 2001; Steffens & Plewe, 2001; Lane et al., 2007). However, the exemplars employed in this experiment were widely used and frequently encountered descriptive words for Gay people. Thus, one interpretation of the results of Experiment 1 could be that the words were either too familiar (i.e., possessing a wide range of very salient stimulus functions) based on their widespread usage in society, or that they held strong negative societal connotations. In either case, the existence of conflicting stimulus functions would have ensured that the 'Gay' functions of the words would not have been the dominant functions of the stimuli employed. Thus, it is possible that while the words may have been semantically descriptive of the target category, conflicting stimulus functions might have ensured they did not instantiate the construct in the required way.

The second experiment in Chapter 2 aimed to build on the findings of Experiment 1 and investigate the relative suitability of multi-exemplar pictorial stimuli. It was conjectured based on previous research that images may in fact be more appropriate for a Homopositivity FAST given that they may elicit more emotional response functions more suited to this particular construct (e.g., Gschwendner et al., 2008; Huijding & de Jong, 2005). Thus, Experiment 2 was identical to the previous experiment, with the only differences being that pictorial rather than verbal stimuli were used. Additionally, a similar known-groups validation technique as in the previous experiment was employed. However, similarly to the first experiment, Experiment 2 showed no predicted effects, with no expected pattern or trend in the data set. Indeed, there was a surprising trend in the data toward strong pre-existing Gay-Good stimulus classes, particularly for the strongly homonegative group.

Thus, it can be assumed that this pictorial FAST in Experiment 2 was an ineffective measure of Homopositivity, given that it failed to generate any differences (based on MHS scores) in learning rates across test blocks. It is likely that this perplexing data pattern was the result of the exemplars being widely known Gay celebrities, such that the images may have been categorised as belonging to a likeable subtype (well-known Gay celebrities) rather than to the class as a whole, thus resulting in confounding cross-class evaluations (see Bluemke & Friese, 2005; De Houwer, 2001). In other words, the positive stimulus functions that may be elicited when presented with images of celebrities may have influenced participants' evaluations of the category as a whole. Similarly, a possible lack of familiarity with the celebrities employed here, or more specifically a lack of familiarity with their sexual orientation, may have resulted in an evaluation of the images in terms of extraneous or irrelevant response functions (e.g., as ‘a White male’, for example).

Another potential confound in Experiment 2 could have been that the exemplars used to represent ‘Gay’ and ‘Good’ did not have strong enough arousing or emotive response

functions that would be required for this particular construct (i.e., homopositive bias). That is, it may have been that exemplars employed were still not evocative enough to engender measurable behavioural resistance in class formation. While the images were selected based on their perceived strong visceral or evocative response functions relative to the previous verbal stimuli, it is possible that they merely elicited relatively neutral response functions that are only distally related to the construct (e.g., 'a Gay person'), rather than of the construct as a whole (i.e., same-sex romantic or sexual behaviour). In a similar vein, it may be possible that the exemplars for the 'Good' category (positive images of people skiing) were not effective exemplars, given that their face validity as strongly positive, evocative stimuli is slightly ambiguous.

Thus, it is possible that either stimulus ambiguity as a result of cross-class confounds or indeed a lack of affective valence may have resulted in the FAST scores being unrelated to MHS scores in Experiment 2. Consequently, based on the findings from the previous chapter, Chapter 3 focused on improving the clarity and affective valence of the images used, and to gain more control over the target stimulus categories. This was primarily to be achieved by employing more generic images of unknown men (i.e., not widely known public figures) to ensure that the target construct is appropriately (and unambiguously) depicted. Chapter 3 also focused on identifying whether images of same-sex couples were more effective in a Homopositivity FAST, given that they should hopefully have eliminated any ambiguity regarding the Gay status of individual exemplars. Thus, Chapter 3 also served to enhance the affective valence of stimuli relative to the previously employed images of solitary individuals by employing romantic and intimate images of same-sex couples.

The first experiment in Chapter 3 was a preliminary investigation into the effectiveness of more emotive and unambiguous pictorial stimuli in a Homopositivity FAST. Essentially, this experiment aimed to obtain in-principle proof that the general FAST

procedure using these exemplars could be a viable tool for assessing attitudes towards sexual minorities. Experiment 3a, therefore, was modified from Experiment 2 by utilising images that were more affective and less ambiguous to represent the target constructs for the Gay-Good stimulus class. In addition, Experiment 3a differed from the previous experiment in that it also incorporated a 'neutral' FAST block alongside the target FAST for comparison purposes, a tactic that has been previously employed in IAT research (Dickson et al., 2013; Jajodia & Earleywine, 2003). This neutral FAST was identical to the Homopositivity FAST with the only exception being that matched pictures of pairs of men in non-intimate situations (e.g., shaking hands) were used in place of the target Gay images. In this way, the primary difference between the two FAST exposures was the romantic or sexual intimacy between the two men, thus allowing for a clear comparison between a pro-man and pro-gay stimulus relation strengths. Upon administering the version of the FAST in Experiment 3a to five participants, an expected trend based on MHS scores was quickly observed (i.e., increases in MHS scores correlated with greater behavioural resistance for the Gay-Good relative to the Men-Good FAST). However, at this point it was realised that it may be possible to improve slightly on the procedure outlined here by replacing the pro-men FAST with an anti-gay FAST. As previously discussed, Experiments 1, 2, and 3a each relied upon finding anti-gay individuals as participants (i.e., a known-groups paradigm) in order to assess the effectiveness of the FAST. That is, it was necessary to secure high numbers of anti-gay individuals against which FAST scores for the rest of the sample could be compared in order to validate the procedure in these experiments. It was decided, therefore, that while the data from Experiment 3a suggested that the procedure was broadly viable, a procedure that does not require strong pre-existing pro or anti-gay bias for validation would be pragmatically beneficial. In other words, it became apparent that what was needed was a paradigm within which a more varied or even a more homopositive participant group could be employed, such

that a test investigating both homopositivity and homonegativity would be a more powerful test as it would allow any population sample to be employed for validation. More specifically, a two-phase Homonegativity/Homopositivity FAST would enable a more comprehensive understanding of the populations' verbal histories related to Gay individuals by providing two opportunities to measure a bias in two different 'directions', much as is achieved by reverse scoring of answers to different questions of the same meaning on a paper and pencil tests. In this way, a more varied population in terms of self-reported Homonegativity could enable validation of the FAST procedure through simple correlational analyses with MHS scores, rather than relying on groups with known anti-gay attitudes.

Experiment 3b, therefore, followed an almost identical format to 3a, with the only exception being that the Men-Good 'neutral' FAST was replaced with a FAST measuring a contrasting Gay-Disgusting affective stimulus relation. It was expected that increases in MHS scores would be generally correlated with greater behavioural resistance for the Homopositivity FAST and with greater behavioural facilitation for the Homonegativity FAST. Indeed, analyses revealed a strong significant correlation between self-reported homonegativity (as ascertained by the MHS) and scores on the FAST. More specifically, this experiment revealed a significant correlation between MHS scores and the Homopositivity FAST, a positive but non-significant correlation with the Homonegativity FAST, and a large significant correlation with the Combined Strength of Relation Indices. Thus, considering that effects are observed within the Homopositivity FAST that were not found in the previous experiments in this thesis, it may be cautiously inferred that the stimuli used in 3b were appropriate at depicting the target constructs for a two-phase Homonegativity/Homopositivity FAST. That is, they may have been useful in detecting participants' verbal histories of stimulus relations. Therefore, based on the compiled evidence from Chapters 2 and 3, it would appear that affective pictorial exemplars might be effective stimuli for a

Homopositivity/Homonegativity FAST, with particular reference to the Homopositivity FAST images.

This would lend support, therefore, to the idea that affective pictorial exemplars are more suitable for a construct such as anti or pro-gay biases than more semantically representative verbal or pictorial stimuli. That is, that they may contain more powerful emotive response functions that better represent the stimulus class of interest than verbal exemplars. As discussed previously, research has suggested that anti-gay attitudes have a strong evaluative component (Hudepohl et al., 2010; Kimmel, 1997; Mahaffey et al., 2005; Parrott, 2008). Additionally, it has been suggested that pictorial exemplars may be more perceptual than words (Henson et al., 2002; Ito & Ulrand, 2005; Joyce et al., 2003; Todorov et al., 2007) and thus more suited to such a construct (Gschwendner et al., 2008). Consequently, this outcome is not necessarily surprising. Interestingly, however, these results contradict the previously discussed IAT research, which typically find word-IATs to be more effective than picture-IATs, in terms of effect sizes (Hurtado, et al., 2009; Mitchell et al., 2003; Nosek et al., 2002), although the current effect may well be confined to tests relating to emotive topics like anti-gay bias.

4.1.2 Increased Stability of a Two-Phase FAST

In addition to the identification of an appropriate and representative stimulus set, another important finding in this thesis was the suggestion in Experiment 3b that the two-phase FAST might be a more stable representation of an individual's verbal history than the standard one-phase FAST. It was conjectured after Experiment 3a that a score that a two-phase FAST examining both the positive and negative attitudes toward a single construct would be a more powerful test. As such, it was also conjectured that a 'combined' score that represented the two FASTs might be a more stable score than the SoR indices for a single FAST independently, such that it would quantify the combination of behavioural resistance

encountered in one FAST with the behavioural facilitation within the other. Indeed, the findings appeared to support this assumption, with significantly stronger effects and correlations observed for the Combined SoR indices relative to the two FASTs independently. Thus, even though each score may provide an absolute measure of an individual's Gay-Good and Gay-Disgusting verbal histories, it would appear that, when combined, they each serve to further inform the other. Simply put, the Combined SoR enables a simple but valid calculation of an individual's behavioural resistance to the formation of a functional Gay-Good stimulus class while accounting for their behavioural facilitation to the formation of an orthogonal stimulus class (i.e., Gay-Disgusting). Consequently, this Combined score may enable a more comprehensive and conceptually valid understanding of an individual's verbal histories related to a particular single construct.

4.1.3 Levels of Bias in an Irish Sample

Another significant and unexpected outcome of this research was the absence of any strong attitudinal bias within the population. Chapter 3 revealed that despite a strong correlation with MHS scores, the sample used in Experiment 3b reported no particular homopositive or homonegative bias at the group level. That is, there was no significant behavioural resistance found to the formation of either a Gay-Good or a Gay-Disgusting stimulus class, suggesting no significantly strong pre-existing pro or anti-gay verbal histories for this sample. This outcome was surprising, given LGB discrimination statistics in Ireland (e.g., O'Higgans-Norman, 2010). Of course, the sample in Experiment 3b was comprised of University-educated individuals under the age of 36, and thus may not be representative of the wider population. Moreover, several individuals showed both MHS and FAST scores indicative of bias, illustrating that bias may still be strong for some individuals. However, given that implicit homonegativity can be quite distinct from self-reported levels (see Cullen & Barnes-Holmes, 2009), it is somewhat unexpected that the trend was not in a slight

homonegative direction for the group as a whole. While this matter undoubtedly needs more extensive and direct investigation, the possibility should of course be considered that there may be a shift occurring away from more traditional (i.e., more explicitly negative or disgust-based) homonegative attitudes in Irish society. Some suggestions for how this might be addressed in future research will be discussed at length later in this Chapter.

4.2 Theoretical and Practical Implications

4.2.1 Implicit Testing

The FAST procedure outlined here undoubtedly has significant theoretical implications for the wider field of implicit testing. Firstly, by attempting to establish functional response classes rather than measuring derived relations (or 'associations' in the social-cognitive literature) the FAST represents the first functionally understood "implicit" test methodology capable of ascertaining the pre-existing strength of a single real-world relation in isolation (e.g., Gay-Good), rather than relative to an opposing construct. Further to this, the novel two-phase pro/anti approach here will enable any researcher interested in a particular construct to obtain a more stable absolute measure of an individual's verbal history from both a positive and negative perspective than could be obtained using a preference-based approach. Additionally, unlike existing measures (e.g., the IAT) the procedure, presentation format, and scoring practices of the FAST were built on fully testable and empirically supported assumptions. Thus, the Homopositivity/Homonegativity FAST developed in this thesis proves further empirical support for the idea that it is indeed possible to identify an individual's histories of relational responding using a test methodology built entirely on functionally understood principles.

Secondly, the current measure is one of relatively few "implicit" measures of homonegativity (see Cullen & Barnes-Holmes, 2008, for a review) and, as such, will contribute to the relatively small field of implicit testing research dedicated to this domain,

especially within an Irish context (see Morrison, Harrington & McDermott, 2010) . Specifically, the two-phase Homopositivity/Homonegativity FAST will address a gap in the literature for a functionally understood and procedurally implicit measure of anti-gay verbal histories, for which only one unpublished study could be identified (Cullen, Barnes-Holmes, Barnes-Holmes, & Stewart, 2007, in Cullen & Barnes-Holmes, 2008). Thus, the current research findings should undoubtedly serve to further our understanding of homonegative prejudice in Ireland, in addition to being of particular interest to behaviour-analytic and social-cognitive researchers interested in the development of more valid and reliable measures of homonegativity more generally.

4.2.2 Practical Applications

In addition to the contributions made by this research to the theoretical field of implicit testing, this research may also be useful for applied researchers and practitioners working within the context of LGB discrimination. Though further research is of course required to stabilise and validate the current FAST procedure, this measure has put forward a functionally understood psychometric-style measure capable of quickly and effectively assessing levels of homonegative prejudice. Furthermore, as stated in Chapter 1, the FAST is one of fastest "implicit" tests currently available within the implicit testing literature, given that it requires only between 5 and 15 minutes to complete. As such, it would surely be of considerable pragmatic benefit to those interested in quickly administering a measure of homonegativity to large populations, such as within a school or work environment.

4.3 Limitations and Issues for Consideration

4.3.1 Practice Effects

Within the current thesis there are of course certain limitations and issues for consideration. The first limitation relates to the practice effects that were observed over the course of the four experiments. Across both experimental chapters, a significant reduction

was consistently observed in the number of trials required to reach the performance criteria from the first to the second baseline blocks, suggesting that participants' performance significantly improved as they progressed through the procedure. As discussed above, research has indicated that practice effects are a pervasive issue within the wider field of implicit testing, with the average test-retest reliability for the widely accepted IAT documented at .56 (Nosek et al., 2007). Within the current thesis, however, addressing the practice effects as comprehensively as is required would have detracted from the main purpose of creating an in-principle Homopositivity FAST prototype. However, there could undoubtedly be implications based on the practice effects across baselines that could undermine the integrity of FAST procedure and, as such, will need to be comprehensively addressed in subsequent research. How this might be adequately addressed in future experiments will be considered later in this discussion.

4.3.2 The Floor Effect

As discussed above, it may be possible that the FAST may not lend itself to the assessment of behavioural facilitation as easily as it does for behavioural resistance. Experiment 3b in the current thesis documented significant correlations between MHS scores and large, negative scores on the Homopositivity FAST, while in contrast, MHS scores did not correlate significantly with large, positive scores on the Homonegativity FAST. In other words, this experiment revealed that the level of behavioural resistance was more pronounced in terms of FAST scores than the corresponding behavioural facilitation. While it is possible given the complexities of human verbal behaviour that this may have been the result of contrasting verbal histories not being bi-directional (i.e., a strong pre-existing Gay-Good stimulus class is not the exact inverse of a weak pre-existing Gay-Disgusting stimulus class), another interpretation could be that it relates to a methodological issue whereby the FAST can more clearly measure resistance relative to facilitation. Simply put, this finding may have

been due to the fact that within the current FAST procedure, if baseline acquisition rates are already relatively low, behavioural facilitation will not be as evident as resistance, given that it is not possible for high fluency behaviours to further increase in fluency to a great degree. Thus, as discussed, it could be possible that the FAST is subject to so-called "floor effects" (Everitt, 2002), whereby behavioural facilitation is intrinsically less detectable than behavioural resistance.

As previously mentioned, it is unlikely that floor effects significantly diminished or negatively affected the outcomes of experiments reported here. This is largely due to the fact that the inclusion of a second Homonegativity block was included merely as a way to provide a more stable and valid combined FAST score, rather than as a requirement for its validation. In other words, the correlation between the MHS and the Homopositivity FAST was significant and informative independent of the outcome of the Homonegativity FAST, which may have suffered from floor effects in the data. Nevertheless, floor effects remain a significant issue for the integrity and validity of the FAST as an implicit measure, and as such will need to be comprehensively addressed in subsequent research. How this might be tackled in future experiments will be discussed in more detail below.

4.3.3 The Validity of the MHS

Another important issue that requires addressing relates to the validity and reliability of the MHS questionnaire. Throughout this thesis, the MHS has been utilised as a general guide for explicit levels of homonegativity, and thus has been employed as a rough indicator of an individual's pro and anti-gay verbal histories. This technique is not uncommon in the implicit testing literature, with explicit scales often used to guide or even validate the outcomes from implicit measures (e.g., Banse et al., 2001; Cunningham, Preacher, & Banaji, 2001; Petróczi, Aidman, & Nepusz, 2008; Skinner, Blick, Coffin, Dudgeon, Forrest, & Morrison, 2013). However, it may seem somewhat paradoxical employing a questionnaire

for validation purposes, given that the issues inherent in self-report measures (i.e., social desirability bias) underpin the entire rationale for the development of an alternative test format for measuring prejudice. It is important to note here, however, that the MHS was employed in this research as a pragmatic guide of explicit homonegativity rather than as a clear indicator of homonegative beliefs. In other words, it was relied upon as a means of providing some knowledge of individual's pre-existing homonegativity but not taken to provide each participant's conclusive homonegativity 'score' against which FAST scores could be compared. However, the fact that the MHS is a measure of so-called 'modern' homonegativity (Morrison & Morrison, 2002), and is comprised of relatively more subtle and indirect questions relating to Homosexuality (e.g., *"*Gay men *seem to focus on the ways in which they differ from heterosexuals and ignore the ways in which they are the same."*), may have ensured its relevance in the current research context. That is, it may have been more suitable than the alternative 'old-fashioned' homonegativity questionnaires available, such as the Attitudes toward Gay and Lesbian Women Scale (Herek, 1988, 1994), given that the purpose of the measure is not as immediately discernible to the participant (Morrison & Morrison, 2002). Based on this, it is not entirely unfeasible that the MHS and the FAST should have been expected to correlate, particularly with regard to the higher scores. Indeed, a general trend based on the MHS was in fact what was aimed for, rather than relying on the MHS as a predictor variable of FAST performance. One important corollary of all of this is that the extent to which the MHS and FAST scores fail to correlate is not a direct reflection on the validity of the FAST, but a reflection of the various degree of validity of the FAST and MHS scores combined. In other words, research in this field is always limited by the validity of previously developed measures.

4.3.4 Influence of Variables such as Gender and Age

It may also be considered a potential limitation that the current research chose not to investigate the relative contributions of variables such as gender and age on FAST scores. That is, any potential differences in the data set according to these variables were not analysed directly. However, the rationale for this reflects a large body of research suggesting that a variety of social and inter-personal factors may contribute to the formation and maintenance of prejudicial attitudes (e.g., Farley, 2000). With regard to homonegativity specifically, there are several factors that have been found to mediate these attitudes, such as age (Herek, 1988; Vuković & Štulhofer, 2012), gender (Gormley & Lopez, 2010; Nagoshi, Adams, Terrell, Hill, Brzuzy, & Nagoshi, 2008), socio-economic status (Hadler, 2012), level of education (Cerny & Polyson, 1984 Elchardus & Spruyt, 2009; Ohlander et al., 2005; Spruyt, 2009), religiosity (Finlay & Walther, 2003; Fulton, Gorsuch & Maynard, 1999; Herek, 1994, Johnson, Brems, & Alford-Keating, 1997), in addition to personality factors such as right-wing conservatism (Cramer, Miller, Amacker, & Burks, 2013), hyper-masculinity (Parrott, Adams, & Zeichner, 2002) and authoritarianism (Warriner, Nagoshi, & Nagoshi, 2013). Similarly, it is possible that additional factors such as, for example, having had direct experience with the LGB community (Besen & Zicklin, 2007; Herek, 2007; Lemm, 2006; Overby & Barth, 2002) or having a Gay family member may mediate homonegative attitudes (Dasgupta & Rivera, 2008). Thus, while it may seem like an oversight to have neglected to analyse variables such as age and gender in the research design, this decision was in fact intentional insofar as it reduced the sample sizes needed and increased the power of planned comparisons conducted during data analysis. Therefore, it should be noted that attempting to examine each of these factors would have required a more complex independent investigation, which would have been outside the scope and intention of this research programme. However, once the basic effect observed here has been validated

with further research, future experiments should of course examine the relative contributions to these variables directly.

4.3.5 Behaviour-Analytic Definition of Attitudes

An additional issue for consideration in this thesis relates to the concept of an attitude and how it is operationally defined within the field of Behaviour Analysis, which was not deeply explored here. This issue is complex within the behaviour-analytic research literature, and is deserving of a comprehensive discussion and elaboration that is beyond the scope of this thesis. However, given that this research is concerned with developing a test format that is widely considered to be appropriate for the measurement of social attitudes and prejudice it is fitting to briefly address the issue here.

The 'attitude' concept is a uniquely social-cognitive construction that is used to conceptualise evaluative cognitive associations that have been established in the experiential history of an individual towards any given element of their environment, which predispose favourable or unfavourable behavioural responses toward that element (see Olsen & Zanna, 1993, for a review). Within Behaviour Analysis, however, the term 'attitude' becomes problematic, in that it represents a hypothetical or mentalistic concept used exclusively for mediative purposes. Within Behaviour Analysis generally, appealing to mediating mental representations or mentalistic terminology as a means to understand and explain human behaviour and language is widely rejected. That being said, as discussed in Chapter 1, behavioural researchers have suggested that it may be possible to conceive attitudes in behaviour-analytic terms as verbal relations between stimuli, and the transformation of the functions of those stimuli in accordance with the network (Hughes et al., 2011).

Within this definition, however, it is not clear what role is played in these networks for non-verbal relations, functional classes of emotional versus purely verbal stimuli, and which types of verbal relations or networks are easiest to measure using implicit tests

methods. In other words, while we can settle upon any definition of attitudes we like based on conceptual criteria, this does not enlighten us as to exactly what types of relations are assessed using implicit testing methods. While the current measure at least makes some progress in providing the bare bones of a real-world implicit test in behavioural terms with real world practical utility for assessing real world attitudes towards sexual minorities, the above issues urgently require substantiation. However, this could easily be addressed in a laboratory investigation. For instance, future experiments could attempt to assess analog verbal networks that are trained as a series of verbal relations consisting of both derived and directly established relations, with both emotional and non-emotional stimulus functions. The FAST could then be used to measure various aspects of these stimulus networks under different parameters (e.g., response windows). Such research would illuminate whether functional classes of stimuli or derived relations are more easily assessed using the FAST, or which affects test outcomes more when these two class types are simultaneously incongruous with each other.

4.4 Suggestions for Future Research

4.4.1 Procedural Refinement of the FAST

In addition to the recommendation above, future research could expand on the findings of this thesis by focusing on the procedural refinement of the FAST methodology. One of the first procedural issues that requires investigation is the practice effect issue that was observed in the current research. As discussed above, significant improvements in learning rates across baselines was observed throughout this thesis, which may undermine the integrity of the FAST. However, this matter can be addressed using some simple methodological alterations. More specifically, the core of the practice effect issue is that the learning (i.e., acquiring a functional response class) is happing at too fast a rate, even if the nature of the stimuli involved modulate this rate. One way to combat this issue, therefore,

could be to decelerate the rate at which participants achieve fluency. This could be achieved by perhaps decreasing the response window (to a point just above the critical threshold needed to establish learning), or by increasing the criterion to complete a block (i.e., the number of trials required to achieve fluency). Both of these modifications lend themselves to easy experimental investigation. By retarding learning across all blocks, it may well be practice effects across blocks will be less pronounced, although this remains an empirical matter.

Another issue that requires systematic investigation in future research is the possible "floor effect" encountered in the current research. As discussed previously, this thesis appeared to suggest that behavioural facilitation is not as easily measured as resistance using the FAST. Given that a learning task that is consistent with one's verbal history cannot get significantly easier relative to a baseline learning task that is not consistent with one's history, this undoubtedly limits the usefulness of the FAST (or at least one of the two FASTs administered if using a two-FAST procedure). However, future experiments that attempt to retard learning rates for the purpose of reducing practice effects would have the added benefit of increasing trial requirements on both baseline blocks and on consistent (i.e., easy) trial blocks. By moving trial requirements further from the zero point, the scope for improvement in learning rates, given a consistent relational learning history, in greatly improved. Consequently, the likelihood of reaching floor effects in the data would be reduced for FASTs that assess facilitated learning. Once again, however, this matter is empirical and we should use this conceptual reasoning only as a heuristic to guide empirical enquiry.

Another issue that could be explored in future research is the possibility that the FAST effect may be the result of some pre-existing familiarity with target stimuli and not with the cross-category stimulus classes (i.e., N stimuli). In other words, a class of stimuli that are merely recognised, rather than related, could in principle explain the current findings.

Further to this, potential pre-existing cross-class confounds for individual participants (i.e., whereby target stimuli may be previously associated with random N stimuli for reasons to do with the idiosyncratic learning histories of various people) may also influence FAST outcomes. While the results of this thesis suggest that neither of these issues were confounding factors here, the possibility that they may have implications for the general FAST format should be investigated in future research. Some ways this might be achieved include running a second FAST using novel N stimuli, which would serve as a check for idiosyncratic pre-existing associations between the first sample of N stimuli employed. Furthermore, ensuring that respectable sample sizes are used and that that multiple exemplars are employed for each stimulus class may additionally diminish any familiarity confounds in future experiments. Similarly, employing these measures would also ensure that any potential procedural confounds related to the enhanced salience of the target stimuli relative to the N categories would also be addressed.

Additionally, future research should focus on refining the validity and reliability of the Homopositivity/Homonegativity FAST procedure. As mentioned above, one of the most important tests for validity in the development of any attitude measure relates to its ability to predict real-world behaviour. This can be dealt with in future research by using explicit behavioural measures as correlates, rather than further paper and pencil tests. This work has recently begin to be addressed in the field of implicit testing, with a few examples of such research to date examining the relationship between the IAT and real-world behaviour (Banaji, 2001; Fazio & Olson, 2003; Greenwald, Poehlman, Uhlmann & Banaji, 2007; Karpinski & Hilton, 2001; Olson & Fazio, 2004; Ziegert & Hanges, 2006). These analyses typically include some behavioural component to the test, whereby the IAT will be correlated against performance on some behavioural task (e.g., the degree of negative or positive interactions with a Black versus a White experimenter; McConnell & Leibold, 2001). A

similar paradigm could easily be employed to validate the FAST in future experiments. For example, FAST performance could be assessed in terms of its ability to predict overt anti-gay behaviours, such as levels of eye-contact with Gay people compared to non-Gay, or the probability of assisting a Gay person in distress compared to a non Gay person.

Another important validation technique that should be examined in future research relates to the stability of the FAST effect. Firstly, this should be examined in future experiments by assessing the stability of the FAST effects over repeated presentations across time. Another interesting way in which the stability of the current test could be addressed could be to modify the design to include several FASTs in one test. One of the primary findings of the current thesis was that the two-phase FAST was more stable than an individual FAST alone. Therefore, it would be an interesting avenue of exploration to analyse the utility of a 'mega-FAST' in which numerous stimulus classes concerning a single construct (i.e., a four or five-phase FAST) were assessed in one large blocks of quasi-randomly sequenced tests (e.g., Gay-Good, Gay-Bad, Gay- Disgust, Gay-Normal, Gay-Dirty, Gay-Clean, Gay-Happy, Gay-Sad, etc). Such a methodology may potentially provide an increasingly stable FAST score, given that each FAST score would inform and thus stabilise the other.

4.4.2 Heteronormativity FAST

In addition to the procedural refinement of the FAST, future research could also address the absence of any strong disgust-based homonegative bias for the sample in Experiment 3b. As discussed above, this finding may suggest that a shift has occurred in Irish society away from more explicit homonegative attitudes. As such, implicit measures of disgust-based homonegativity may be harder to validate within an Irish context and may indeed be of less need in these domains. However, a review of the current statistics surrounding the legal and societal discrimination against LGB individuals in Ireland would indicate that anti-gay attitudes remain widespread in Irish society and, consequently, it may

be necessary to devote further research to develop a Homonegativity FAST that more accurately targets anti-gay attitudes within an Irish population. To achieve this, it is necessary for future experiments to explore alternative avenues than the disgust-based FAST utilised here. However, a review of the literature would indicate that anti-gay attitudes are particularly complex and multi-faceted, and may reflect a set of widespread socio-cultural stereotypes about sexual minority groups as opposed to a generalised dislike or hatred of the LGB population. For instance, some studies have suggested that anti-effeminacy attitudes may a particularly problematic factors in anti-gay attitudes (Parrott et al., 2002; Taywaditep, 2001). Similarly, factors such as gender role stereotyping (Kite & Whitley, 1996) as well as masculinity and femininity ideals (Cohen, Hall, & Tuttle, 2009; Overby & Barth, 2002; Sánchez, Greenberg, Liu, & Vilain, 2009) have also been argued to be important underlying factors to consider in discriminatory anti-gay attitudes.

Some researchers have suggested that one of the most pressing issues with regard to LGB discrimination, however, is deeply rooted societal hetero*normativity* rather than homonegativity (Oesterreich, 2002; Hegarty, Pratto, & Lemieux, 2004). Heteronormativity has been defined as "the view that institutionalized heterosexuality constitutes the standard for legitimate and expected social and sexual relations" (Ingraham, 1999, p. 17). In contrast to homonegativity, heteronormativity is argued to reflect the set of underlying legal and socio-cultural assumptions that view heterosexuality is the only natural, inevitable, and desirable sexuality (Kitzinger, 2005). Heteronormative values are argued to pervade society in a wide variety of ways, such as through prescriptive gender roles (Eagly, 1987) and normative expectations about heterosexuality (see Hegarty et al., 2004), in addition to the widespread societal privileges associated with having a heterosexual identity (see Massey, 2009). Heteronormative attitudes are proposed to reflect these assumptions, and are defined as the set of affective responses and beliefs that prioritise and value heterosexuality, in

addition to stigmatising all sexual minorities who deviate from this perceived natural social norm (Hegarty et al., 2004). In this way, the legal and socio-cultural means by which heteronormativity is perpetuated inherently limit who is considered a full citizen in society (Blumenfeld, 1992). The absence of marriage and parenting equality for LGB individuals, for example, has been argued to reflect social and legal heteronormative family assumptions that regard heterosexual couples with children as the superior and 'natural' family (DePalma & Atkinson, 2009).

Given the foregoing, future experiments could focus on the development of a FAST for Heteronormativity (i.e., Gay-Normal rather than Gay-Good). This would be beneficial firstly as it is a pressing social issue, and also given that it may be more reflective of modern attitudes toward sexual minority groups in Ireland. Furthermore, the issue of implicit heteronormativity has not as of yet been examined and, consequently, focusing on it directly would address an important gap in the implicit testing literature. Additionally, adopting a heteronormative approach may be more well-suited to examining prejudicial attitudes to lesbian women, which were not investigated here. Previous research has suggested that an important distinction exists between attitudes to Gay men and Lesbian women, with stronger homonegative prejudice and anti-gay aggression typically found toward Gay men (Breen & Karpinski, 2013; Morrison & Morrison, 2011). However, it would be interesting to examine prejudicial attitudes from a heteronormative perspective and investigate whether a similar trend is observed. This would not only address a gap in the literature, but could also provide a more valid and accurate reflection of LGB prejudice and discrimination.

4.5 Concluding Remarks

Over the course of four experiments, the Function Acquisition Speed Test has been refined and developed into a functionally understood prototype for the assessment of homonegativity. More specifically, this research has identified an effective and representative

stimulus set, in addition to establishing a reliable and stable two-phase FAST format capable of ascertaining an individual's homonegative verbal histories. Thus, these findings have added firstly to the growing body of literature on behaviour-analytic implicit measures, as well as providing a useful and easy-to-administer tool for the measurement of homonegative prejudice. Therefore, this thesis has succeeded in its goal of extending the FAST to the arena of social prejudice. In doing so, it has ensured that it is now at least tenable to begin a serious agenda of social research into the concept of attitudes equipped with tools that are functionally understood, even if the attitude concept itself is not. Indeed, to conclude this thesis, it should be noted that it will be through empirical research of this kind that we slowly triangulate in on the behavioural phenomena measured by these tests, and, potentially, get increasingly closer to a functional understanding of the set of behaviours that social researchers refer to as 'attitudes'.

References

Ahmad, S., & Bhugra, D. (2010). Homophobia: a review. *Sex and Relationship Therapy. 25(4),* 447-455.

Aidman, E. V., & Carroll, S. M. (2003). Implicit individual differences: Relationships between implicit self-esteem, gender identity, and gender attitudes. *European Journal of Personality, 17*, 19-36.

Aiken, L. R. (1994). *Psychological Testing and Assessment*. (8th ed.). Massachusetts: Simon & Schuster, Inc.

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173-221). Mahwah, NJ: Erlbaum.

Allen, D. & Oleson, T. (1999). Shame and internalised homophobia in Gay men. *Journal of Homosexuality, 37(3),* 33-43.

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of Social Psychology*. Worcester, Mass: Clark University Press.

Arkes, H., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or "Would Jesse Jackson 'fail' the Implicit Association Test?". *Psychological Inquiry, 15*, 257-278.

Balsam, K. F., Beauchaine, T. P., Mickey, R. M., & Rothblum, E. D. (2005). Mental health of lesbian, Gay, bisexual, and heterosexual siblings: Effects of gender, sexual orientation, and family. *Journal of Abnormal Psychology*. *114(3),* 471-476.

Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger, III, J. S. Nairne, I. Neath, & A. Surprenant (Eds.), The nature of remembering: Essays in honor of Robert G. Crowder (pp. 117-150). Washington, DC: American Psychological Association.

Banks, C. (2003). *The cost of Homophobia: Literature Review on the Human Impact of Homophobia in Canada*. Saskatoon: Gay and Lesbian Health Services, Saskatoon

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes toward homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift fur Experimentelle Psychologie, 48,* 145-160.

Barnes, D., Lawlor, H., Smeets, P. M., & Roche, B. (1996). Stimulus equivalence and academic self-concept in mildly mentally handicapped and non-mentally handicapped children. *The Psychological Record, 46,* 87-107.

Barnes, D., McCullagh, P. D., & Keenan, M. (1990). Equivalence class formation in non-hearing impaired children and hearing impaired children. *The Analysis of Verbal Behavior, 8,* 19-30.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *The Psychological Record, 60,* 527-542.

Barnes-Holmes, D., Hayden, E., Barnes-Holmes, Y., & Stewart, I. (2008). The Implicit Relational Assessment Procedure (IRAP) as a response-time and event-related-potentials methodology for testing natural verbal relations: A preliminary study. *The Psychological Record, 58,* 497-516

Barnes-Holmes, D., Murphy, A., Barnes-Holmes, Y., & Stewart, I. (2010). The Implicit Relational Assessment Procedure (IRAP): Exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-Black stereotyping among white Irish individuals. *The Psychological Record, 60,* 57-66.

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6, 10 & adulthood. *Psychological Science, 17,* 53-58.

Berrill, K. T. (1992). AntiGay violence and victimization in the United States: An overview. In G. M., Herek, & K. Berrill, (Eds.). *Hate crimes: Confronting violence against lesbians and Gay men* (pp. 19-35). Thousand Oaks, CA: Sage.

Besen, Y., & Zicklin, G. (2007). Young men, religion and attitudes towards homosexuality. *Journal of Men, Masculinities and Spirituality, 1 (3),* 250-266.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6,* 242-261.

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist, 61,* 27-41.

Blanton, H, & Jaccard, J. (2008). Unconscious racism: A concept in pursuit of a measure. *Annual Review of Sociology, 34,* 277-297.

Blanton, H., Jaccard, J., Gonzales, P. M. & Christie, C. (2006). Decoding the Implicit Association Test: Implications for criterion prediction. *Journal of Experimental Social Psychology,* 42, 192-212.

Bluemke, M., & Friese, M. (2006). Do features of stimuli influence IAT effects? *Journal of Experimental Social Psychology, 42(2),* 163-176.

Blumenfeld, W. J. (1992). *Homophobia: How we all Pay the Price*. Boston, MA: Beacon.

Breen, A. B., & Karpinski, A. (2013). Implicit and explicit attitudes toward Gay males and lesbians among heterosexual males and females. *Journal of Social Psychology, 153(3),* 351-374.

Brevers, D., Cleeremans, A., Hermant, M. A., Tibboel, H., Kornreich, C., Verbanck, P., & Noël, X. (2013). Implicit gambling attitudes in problem gamblers: Positive but not negative implicit associations. *Journal of Behavior Therapy and Experimental Psychiatry, 44,* 94-97.

Brunel, F. F., Tietje, B. C., & Greenwald, A. G. (2004). Is the Implicit Association Test a valid and valuable measure of implicit consumer social cognition. *Journal of Consumer Psychology, 14,* 385–404.

Cerny, J. A. & Polyson, J. (1984). Changing homonegative attitudes. *Journal of Social and Clinical Psychology, 2(4),* 366-371.

Chiu, P., Ambady, N., & Deldin, P. J. (2004). Contingent Negative Variation to emotional in- and out-group stimuli differentiates high- and low- prejudiced individuals. *Journal of Cognitive Neuroscience, 16(10),* 1830-1839.

Cochran, S. D., (2001). Emerging issues in research on lesbians' and Gay men's mental health: does sexual orientation really matter? *The American Psychologist, 56(11),* 931-947.

Cohen, T. R., Hall, D. L., & Tuttle, J. (2009). Attitudes toward stereotypical versus counterstereotypical Gay men and lesbians. *Journal of Sex Research, 46,* 1-8.

Coker T. R., Austin S. B., & Schuster M. A. (2010). The health and health care of Gay, lesbian, and bisexual youth. *Annual Review of Public Health, 31,* 457-477.

Collier, G., Minton, H. L., & Reynolds, G. (1991). *Currents of Thought in American Social Psychology* . New York: Oxford.

Cottrell, C. A., & Neuberg, S. L. (2005). Different emotional reactions to different groups: A sociofunctional threat-based approach to 'prejudice'. *Journal of Personality & Social Psychology, 88,* 770-789.

Cramer, R.J., Miller, A.K., Amacker, A.M., & Burks, A.C. (2013). Openness, Right-wing Authoritarianism, and AntiGay Prejudice in College Students: A Mediation Model. *Journal of Counseling Psychology, 60,* 64-71.

Cronbach, L. J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281-302.

Cullen, C., & Barnes-Holmes, D. (2008). Implicit pride and prejudice: A heterosexual phenomenon? In T. G. Morrison & M. A. Morrison (Eds.), *Modern prejudice* (pp. 195-223). New York: Nova Science.

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

Cvencek, D., Greenwald, A. G., Brown, A. S., Gray, N. S., & Snowden, R. J. (2010). Faking of the Implicit Association Test is statistically detectable and partly correctable. *Basic and Applied Social Psychology, 32,* 302-314.

Czaja, S. J., & Sharit, J. (1993). Age differences in the performance of computer-based work. *Psychology and Aging, 8,* 59-67.

Czopp, A. M. & Monteith, M. J. (2003). Confronting prejudice (literally): Reactions to confrontations of racial agender bias. *Personality and Social Psychology Bulletin, 29,* 532-544.

D'Augelli, A. R., Grossman, A. H., & Starks, M. T. (2006). Childhood gender atypicality, victimization, and PTSD among lesbian, Gay, and bisexual youth. *Journal of Interpersonal Violence, 21,* 1-21.

Dambrun, M., & Guimond, S. (2004). Implicit and explicit measures of prejudice and stereotyping: Do they assess the same underlying knowledge structure? *European Journal of Social Psychology, 34,* 663-676.

Dasgupta, N., & Rivera, L. M. (2008). When social context matters: The influence of long-term contact and short-term exposure to admired outgroup members on implicit attitudes and behavioral intentions. *Social Cognition, 26,* 54-66.

De Houwer, J., & De Bruycker, E. (2007). The Implicit Association Test outperforms the Extrinsic Affective Simon Task as a measure of interindividual differences in attitudes. *British Journal of Social Psychology, 46, 401*-421.

De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. Journal of Experimental Social Psychology, 37, 443- 451.

De Houwer, J. (2002). The implicit association test as a tool for studying dysfunctional associations in psychopathology: Strengths and limitations. *Journal of Behavior Therapy and Experimental Psychiatry, 33,* 115-133.

De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology, 50(2),* 77-85.De Houwer, 2006

De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science, 6,* 202-209.

De Jong, M.G., Pieters, F.G.M. & Stremersch, S. (2012). Analysis of sensitive questions across cultures: An application of multigroup item randomized response theory to sexual attitudes and behavior. *Journal of Personality and Social Psychology, 103(3), 5*43-564.

DePalma, R., & Atkinson, E. (2009). "Permission to talk about it": narratives of sexualities equality in the primary classroom. *Qualitative Inquiry 15(7),* 876-892.

Devaney, J. M., Hayes, S. C., & Nelson, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior, 46,* 243-257.

Devos, T., & Ma, D. (2008). Is Kate Winslet more American than Lucy Liu? The impact of construal processes on the implicit ascription of a national identity. *British Journal of Social Psychology, 47,* 191-215.

Dickson, J. M., Gately, C., & Field, M. (2013) Alcohol dependence is characterised by weak negative rather than strong positive implicit alcohol associations. *Psychopharmacology, 228,* 603-610.

Dixon, M. R., Rehfeldt, R. A., Zlomke, K. R., & Robinson, A. (2006). Exploring the development and dismantling of equivalence classes involving terrorist stimuli. *Psychological Record, 56,* 83-103.

Donders, F. C. (1969). On the speed of psychological processes. In W. Koster (Ed.), *Attention and Performance: II* (Original work published 1868). Amsterdam: North-Holland.

Drake, C. E., Kellum, K. K., Wilson, K. G., Luoma, J. B., Weinstein, J., & Adams, C. H. (2010). Examining the Implicit Relational Assessment Procedure: Four Pilot Studies. *Psychological Record, 60, 81-*86.

Dymond, S., & Whelan, R. (2010). Derived relational responding: A comparison of matching to sample and the relational completion procedure. *Journal of the Experimental Analysis of Behavior, 94,* 37-55.

Dymond, S. & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more-than, and less-than. *Journal of the Experimental Analysis of Behavior, 64,* 163-184. Erratum, *66,* 348.

Eagly, A. H. (1987). *Sex Differences in Social Behavior: A Social-Role Interpretation*. Hillsdale, NJ: Erlbaum.

Egloff, B. & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology, 83,* 1441-1455.

Egloff, B., Wilhelm, F. H., Neubauer, D. H., Mauss, I. B., & Gross, J. J. (2002). Implicit anxiety measure predicts cardiovascular reactivity to an evaluated speaking task. *Emotion, 2,* 3-11.

Everitt, B. S. (2002). *Cambridge Dictionary of Statistics* (2nd Ed). Cambridge University Press.

Farley, J. E. (2000). *Majority - Minority Relations.* (4th Ed.) Upper Saddle River, NJ: Prentice Hall.

Fazio, R. H., & Olson, M. A. (2003). Attitudes: Foundations, functions, and consequences. In M. A. Hogg & J. Cooper (Eds.), *The handbook of social psychology* (pp 139-160). London: Sage.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp 247-282). Hillsdale, NJ: Erlbaum.

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

Fenton, K. A., Johnson, A. M., McManus, S., & Erens, B. (2001). Measuring sexual behaviour: Methodological challenges in survey research. *Sexually Transmitted Infections, 77,* 84-92.

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

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the 'I', the 'A', and the 'T': A logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology, 17,* 74-147.

Finlay, B., & Walther, C. S. (2003). The relation of religious affiliation, service attendance, and other factors to homophobic attitudes among university students. *Review of Religious Research, 44,* 370 – 393.

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

Foroni, F. & Bel-Bahar, T. S. (2010). Picture-IAT vs. Word-IAT: Level of stimulus representation influences on the IAT. *European Journal of Social Psychology, 40,* 321-337.

Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The race Implicit Association Test as a stereotype threat experience. *Personality and Social Psychology Bulletin, 30,* 1611-1624.

Friedman, R. C. (1999). Homosexuality, psychopathology, and suicidality. *Archives of General Psychiatry, 56,* 887-888.

Fulton, A. S., R. L. Gorsuch, and E. A. Maynard. 1999. Religious orientation, antihomosexual sentiment, and fundamentalism among Christians. *Journal for the Scientific Study of Religion, 38(1),* 14–22.

Fumham, A. (1986). Response bias, social desirability, and dissimulation. *Personality and Individual Differences, 7,* 385-400.

Gavin, A., Roche, B., & Ruiz, M. R. (2008). Competing contingencies over derived relational responding: A behavioral model of the Implicit Association Test. *The Psychological Record, 58,* 427-441.

Gavin, A., Roche, B., Ruiz, M. R. & Hogan, M., & O'Reilly (2012). A behavior-analytically modified Implicit Association Test for measuring the sexual categorization of children. *The Psychological Record, 62,* 55-68.

Gawronski, B., & Payne, B.K., (Eds.). (2010). *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications.* New York, NY: Guilford Press.

Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology, 50,* 141-150.

Glashouwer, K. A., Smulders, F. T. Y., de Jong, P. J., Roefs, A., & Wiers, R. W. (2013). Measuring automatic associations: a comparison of algorithms for the Implicit Association Test (IAT) in a laboratory setting. *Journal of Behavior Therapy and Experimental Psychiatry, 44,* 105-113.

Gormley, B., & Lopez, F. G. (2010). Correlates of psychological abuse perpetration in college dating relationships. *Journal of College Counseling, 13,* 4-16.

Govan, C. L., & Williams, K. D. (2004). Reversing or eliminating IAT effects by changing the affective valence of the stimulus items. *Journal of Experimental Social Psychology, 40(3),* 357-365.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102(1),* 4-27.

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79,* 1022-1038.

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., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,*197-216.

Greenwald, A. G., Oakes, M. A., & Hoffman, H. (2003). Targets of Discrimination: Effects of Race on Responses to Weapons Holders. *Journal of Experimental Social Psychology, 39,* 399-405.

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

Grey, I., & Barnes, D. (1996). Stimulus equivalence and attitudes. *The Psychological Record, 46,* 243-270.

Gschwendner, T., Hofmann, W., & Schmitt, M. (2008). Convergent and predictive validity of implicit and explicit anxiety measures as a function of specificity similarity and content similarity. *European Journal of Psychological Assessment, 24,* 254-262.

Haddock , G. Zanna , M. P. (1993). Assessing the structure of prejudicial attitudes: The case of attitudes toward homosexuals. *Journal of Personality and Social Psychology, 65(6),* 105-1118.

Hadler, M. (2012). The influence of world societal forces on social tolerance. A time comparative study of prejudices in 32 countries. In: *The Sociological Quarterly, 53(1),* 211-237.

Haidt, J., Rozin, P., McCauley, C., & Imada, S. (1997). Body, psyche, and culture: The relationship of disgust to morality. *Psychology and Developing Societies, 9,* 107-131.

Hall, G., Mitchell, C., Graham, S., & Lavis, Y. (2003). Acquired equivalence and distinctiveness in human discrimination learning: Evidence for associative mediation. *Journal of Experimental Psychology: General, 132(2),* 266–276.

Hayes, S. C., & Wilson, K. G. (1993). Some applied implications of a contemporary behavior analytic account of verbal behavior. *The Behavior Analyst, 16,* 283-301.

Hayes, S. C. (1994). Content, context, and the types of psychological acceptance. In S. C. Hayes, N. S. Jacobson, V M. Follette, & M. J. Dougher (Eds.), *Acceptance and change. content and context in psychotherapy* (pp. 125-140). Reno, NV: Context Press.

Hegarty, P. (2002). It's not a choice, it's the way we're built: Symbolic beliefs about sexual orientation in the US and Britain. *Journal of Community and Applied Social Psychology, 12,* 153–166.

Hegarty, P., Pratto, F., & Lemieux, A. F. (2004). Heterosexist ambivalence and heterocentric norms: Drinking in intergroup discomfort. *Group Processes and Intergroup Relations, 7,* 119-130.

Heider, J. D. & Skowronski, J.J. (2007). Improving the Predictive Validity of the Implicit Association Test. *North American Journal of Psychology, 9(1),* 53-76.

Henson, R.N.A., Price, C.J., Rugg, M.D., Turner, R., & Friston, K. J. (2002). Detecting latency differences in event-related BOLD responses: application to words versus non-words and initial versus repeated face presentations. *NeuroImage, 15(1),* 83-97.

Herek, G.M. (1988). Heterosexuals' attitudes toward lesbians and Gay men: Correlates and gender differences. *Journal of Sex Research, 25,* 451-477.

Herek, G.M. (1994). Assessing heterosexuals' attitudes toward lesbians and Gay men: A review of empirical research with the ATLG scale. In B. Greene, & G.M. Herek (Eds.) *Lesbian and Gay psychology: Theory, research, and clinical applications* (pp. 206-228). Thousand Oaks, CA: Sage Publications.

Herek, G.M. (2000). Sexual prejudice and gender: Do heterosexuals' attitudes toward lesbians and Gay men differ? *Journal of Social Issues 56(2),* 251-266.

Herek, G. M. (2007). Confronting sexual stigma and prejudice: Theory and practice. *Journal of Social Issues, 63 (4),* 905-925.

Herek, G. M., Cogan, J. C., Gillis, J. R., & Glunt, E. K. (1998). Correlates of internalized homophobia in a community sample of lesbians and Gay men. *Journal of the Gay and Lesbian Medical Association, 2,* 17-25.

Hicks, G. R., & Lee, T. (2006). Public attitudes toward Gays and lesbians: Trends and predictors. *Journal of Homosexuality, 51,* 57- 77

Hodson, G., & Costello, K. (2007). Interpersonal disgust, ideological orientations, and dehumanization as predictors of intergroup attitudes. *Psychological Science, 18,* 691-698.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2004). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. Unpublished manuscript, University of Trier, Germany.

Holbrook, A. L., & Krosnick, J. A. (2010). Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly, 74,* 37-67.

Hudepohl, A. D., Parrott, D. J., & Zeichner, A. (2010). Heterosexual men's anger in response to male homosexuality: Effects of erotic and non-erotic depictions of male-male intimacy and sexual prejudice. *Journal of Homosexuality, 57(8),* 1022-1038.

Hudson, W. W., & Ricketts, W. A. (1980). A strategy for the measure- ment of homophobia. *Journal of Homosexuality, 5,* 356-371.

Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorising in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record, 61,* 465–498.

Huijding, J., & de Jong, P. J. (2005). A pictorial version of the Extrinsic Affective Simon Task: Sensitivity to generally affective and phobia-relevant stimuli in high and low spider fearful individuals. *Experimental Psychology, 52,* 289–295.

Hurtado, E., Gonzalez, R., Haye, A., Manes, F. & Ibanez, A. (2009). Contextual blending of ingroup/outgroup face stimuli and word valence: LPP modulation and convergence of measures. *BMC Neuroscience, 101(69),* 1-21.

Inbar, Y., Pizarro, D. A., Knobe, J., & Bloom, P. (2009). Disgust sensitivity predicts intuitive disapproval of Gays. *Emotion, 9,* 435-439.

Ingraham, C. (1999). *White Weddings: Romancing Heterosexuality in Popular Culture.* New York: Routledge.

Itaborahy, L. P., & Zhu, J. (2013). State-sponsored Homophobia: A world survey of laws: Criminalisation, protection and recognition of same-sex love. The International Lesbian, Gay, Bisexual, Trans and Intersex Association.

Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, and Behavioral. Neuroscience, 5,* 21-36

Jajodia, A., & Earleywine, M. (2003). Measuring alcohol expectancies with the Implicit Association Test. *Psychology of Addictive Behaviors, 17,* 126−133.

Jellison, W. A., McConnell, A. R., & Gabriel, S. (2004). Implicit and explicit measures of sexual orientation attitudes: Ingroup preferences and related behaviors and beliefs among Gay and straight men. *Personality and Social Psychology Bulletin, 30,* 629-642.

Johnson, M. E., Brems, C., & Alford-Keating, P. (1997). Personality correlates of homophobia. J*ournal of Homosexuality, 34(1),* 57-69.

Jones, K., & Devos, T. (2013). Gay men's implicit attitudes towards sexual orientation: Disentangling the role of sociocultural influences and social identification. *Psychology and Sexuality,* 1-17

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology, 25,* 881-919.

Jost, J. T., Pelham, B. W., & Carvallo, M. (2002). Non-conscious forms of system justification: Cognitive, affective, and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology, 38,* 586-602.

Kang, J., & Banaji, M. R. (2006). Fair measures: A behavioral realist revision of "affirmative action.". *California Law Review, 94,* 1063-1118.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81,* 774–778.

Kehily, M. J. (2001) Understanding heterosexualities: masculinities, embodiment and schooling. *Men and Masculinities, 4(2),* 173-185.

Kim, D. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly, 66, 1,* 83-96.

Kimmel, M. S. (1997). Masculinity as homophobia: Fear, shame and silence in the construction of gender identity. In M.M. Gergen & S.N. Davis (Eds.), *Toward a new psychology of gender* (pp. 223-242). New York. Routledge.

Kitzinger, C. (2005). Heteronormativity in action: Reproducing the heterosexual nuclear family in after-hours medical calls. *Social Problems, 52(4),* 477-498.

Klassen, D., Anderson, P. B., & Hornstra, R. K. (1975). Influence of social desirability on symptom and mood reporting in a community survey. *Journal of Consulting and Clinical Psychology, 43,* 448-452.

Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order eVects: New evidence for task-set switching account for the Implicit Association Test eVects. *Personality and Social Psychology Bulletin, 31,* 208–217.

Kruglanski, A. W., & Stroebe, W. (Eds.) (2012). *Handbook of the History of Social Psychology.* New York: Psychology Press.

Lacerda, M., Pereira, C., &. Camino, L (2002). Um estudo sobre as formas de preconceito contra os homossexuais na perspectiva das representações sociais [A study of prejudice forms against homosexuals anchored on social representations]. *Psicologia: Reflexão e Crítica, 15(1),* 165-178.

Lane, K. A., Kang, J., & Banaji, M. R. (2007). Implicit social cognition and law. *Annual Review of Law and Social Science, 3,* 427-451.

Lang, P. J., Bradley, M. M., & Cuthbert, B. (1995). *International Affective Picture System.* Gainesville: University of Florida, Center for Research in Psychophysiology.

Leigland, S. (1997). Is a new definition of verbal behavior necessary in light of derived relational phenomena? *The Behavior Analyst, 20,* 3-9.

Lemm, K. M. (2006). Positive associations among interpersonal contact, motivation, and implicit and explicit attitudes toward gay men. *Journal of Homosexuality, 51 (9),* 79-99.

Lipkens, G., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology, 56,* 201-239.

Logan, C. (1996) Homophobia? No, homoprejudice. *Journal of Homosexuality, 31,* 31–53.

Mac an Ghaill, Hanafin, J., & Conway, P. F. (2004). *Gender Politics and Exploring Masculinities in Irish Education: Teachers, Materials and the Media.* Dublin: National Council for Curriculum and Assessment.

Mace, F. C., Hock, M. L., Lalli, J. S., West, B. J., Belfiore, P., Pinter, E., & Brown, D. K. (1988). Behavioral momentum in the treatment of noncompliance. *Journal of Applied Behavior Analysis, 21,* 123-141.

Mahaffey, A. L., Bryan, A., & Hutchison, K. E. (2005). Startle eyeblink as a potential implicit measure of homophobia. *Basic and Applied Social Psychology, 27,* 37-45.

Maison, D., Greenwald, A., & Bruin, R. H. (2001). The Implicit Association Test as a measure of implicit consumer attitudes. *Polish Psychological Bulletin, 32,* 61–70.

Marsolek, C. J. (1999). Dissociable neural subsystems underlie abstract and specific object recognition. *Psychological Science, 10,* 111-118.

Massey, S. G. (2009). Polymorphous prejudice: Liberating the measurement of heterosexuals' attitudes toward lesbians and Gay men. *Journal of Homosexuality, 56(2),* 147 – 172.

Maycock, P., Bryan, A., Carr, N. & Kitching, K. (2009). *Supporting LGBT Lives: A Study of the Mental Health and Well-Being of Lesbian, Gay, Bisexual and Transgender People.* Dublin: Gay and Lesbian Equality Network and Belong To Youth Service.

McConnell, A. R., & Liebold, J. M. (2001). Relations between the Implicit Association Test, explicit racial attitudes, and discriminatory behavior. *Journal of Experimental Social Psychology, 37,* 435–442.

McFarland, S., & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition, 20(6),* 483-510.

McGlinchey, A., Keenan, M., & Dillenburger, K. (2000). Outline for the development of a screening procedure for children who have been sexually abused. *Research on Social Work Practice, 10,* 722-747.

McKenna, I. M., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2007). Testing the fake-ability of the Implicit Relational Assessment Procedure (IRAP): The first study. International *Journal of Psychology and Psychological Therapy, 7,* 253-26

Merwin, R. M., & Wilson, K. G. (2005). Preliminary findings on the effects of self-referring and evaluative stimuli on stimulus equivalence class formation. *The Psychological Record, 55,* 561–575.

Meyer, I. & Dean, L. (1998). Internalized homophobia, intimacy and sexual behaviour among Gay and bisexual men. In Herek, G. (ed.), *Stigma and sexual orientation* (pp. 160–186). Sage, Thousand Oaks, CA.

Meyer I. (2003). Prejudice as stress: Conceptual and measurement problems. *American Journal of Public Health, 93,* 262–265.

Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology, 85(6),* 1180-1192.

Millar, M. G., & Tesser, A. (1990). Attitudes and behaviour: The cognitive-affective mismatch hypothesis. *Advances in Consumer Research, 17,* 86-90.

Minton, S. J., Dahl, T., O'Moore, A. M. & Tuck, D. (2008). An exploratory survey of the experiences of homophobic bullying among lesbian, Gay, bisexual, and trasngendered young people in Ireland. *Irish Educational Studies, 27(2),* 177-191.

Mitchell, C. J., Anderson, N. E., & Lovibond, P. F. (2003). Measuring evaluative conditioning using the Implicit Association Test. *Learning and Motivation, 34,* 203–217.

Mitchell, J. A., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132(3),* 455-469.

Mohipp, C., & Morry, M. (2004). The relationship of symbolic beliefs and prior contact to heterosexuals' attitudes toward Gay men and lesbian women. *Canadian Journal of Behavioural Science, 36,* 36–44.

Morgan, Ch.D. & Murray, H.A. (1935). A method for investigating fantasies: The Thematic Apperception Test. Archives of Neurology and Psychiatry, 34, pp. 289-306.

Morrison, M. A., & Morrison, T. G. (2011). Sexual orientation bias toward Gay men and lesbian women: Modern homonegative attitudes and their associations with discriminatory behavioural intentions. *Journal of Applied Social Psychology, 41,* 2573-2599.

Morrison, M. A., & Morrison, T. G. (2002). Development and validation of a scale measuring modern prejudice toward Gay men and lesbian women. *Journal of Homosexuality, 43(2),* 15-37.

Morrison, T. G., Harrington, R., & McDermott, D. T. (2010). Bi now, Gay later: Implicit and explicit binegativity among Irish university students. *Journal of Bisexuality, 10(3),* 211 - 232.

Morrison, T.G., Kenny, P., & Harrington, A. (2005). Modern prejudice toward Gay men and lesbian women: Assessing the viability of a measure of modern homonegative attitudes within an Irish context. *Genetic, Social, and General Psychology Monographs, 131*, 219-250.

Moxon, P. D., Keenan, M., & Hine, L. (1993). Gender-role stereotyping and stimulus equivalence. The Psychological Record, 43, 381–394.

Nagoshi, J. L., Adams, K. A., Terrell, H. K., Hill, E. D., Brzuzy, S., & Nagoshi, C. T. (2008). Gender differences in correlates of homophobia and transphobia. *Sex Roles, 59,* 521-531.

Nelson, D. L., & Schreiber, T. A. (1992). Word concreteness and word structure as independent determinants of recall. *Journal of Memory and Language, 31,* 237-260.

Nettler, G., & Golding, E. H. (1946). The measurement of attitudes toward the Japanese in America. *American Journal of Sociology, 52(1),* 31-40.

Nierman, A. J., Thompson, S. C., Bryan, A., & Mahaffey, A. L. (2007). Gender role beliefs and attitudes toward lesbians and Gay men in Chile and the U.S. *Sex Roles, 57,* 61-67.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35(4),* 250-256.

Norman, J. (2005). A survey of teachers on homophobic bullying in Irish second-level schools. *Dublin: Dublin City University.*

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134*, 565-584.

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(2),* 166-180.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes (pp. 265-292). New York: Psychology Press.

O'Higgins-Norman, J. Galvin, M. & McNamara, G. (2006). Straight Talk: Researching Gay & Lesbian Issues in the School Curriculum. *Dublin: Dublin City University.*

O'Reilly, A., Roche, B., Ruiz, M., Ryan, A., & Campion, G. (In press). A Function Acquisition Speed Test for equivalence relations (FASTER). *The Psychological Record.*

O'Reilly, A., Roche, B., Ruiz, M. R., Tyndall, I. & Gavin, A. (2012). The Function Acquisition Speed Test (FAST): A behavior-analytic implicit test for assessing stimulus relations. *The Psychological Record, 62,* 507–528.

O'Toole, C. & Barnes-Holmes, D. (2007). *Using the IRAP to Investigate the Link between Human Relational Processes and Intelligence.* Paper presented in the symposium; Relational Frame Theory: Clinical and Educational Research, held during the 33rd Annual Convention for the Association of Behavioural Analysis in San Diego, California.

Oesterreich, H. (2002). "Outing" social justice: Transforming civic education within the challenges of heteronormativity, heterosexism and homophobia. *Theory and Research in Social Education 30(2),* 287-301.

Ohlander, J., Batalova, J., & Treas, J. (2005). Explaining educational influences on attitudes toward homosexual relations. *Social Science Research, 34 (4),* 781-799.

Olatunji, B. O. (2008). Disgust, scrupulosity and conservative attitudes about sex: Evidence for a mediational model of homophobia. *Journal of Research in Personality, 42,* 1364-1369.

Olsen, J. M., & Zanna, M. P. (1993). Attitudes and attitude change. *Annual Review of Psychology, 44,* 117-154.

Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science, 14,* 636–639.

Olson, M. A., & Fazio, R. H. (2004). Reducing the Influence of Extrapersonal Associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86,* 653–667.

Ostrom, T. M. (1969). The relationship between the affective, behavioural, and cognitive components off attitudes. *Journal of Experimental Social Psychology, 5,* 12-30.

Overby, L. M., & Barth, J. (2002). Contact, community context, and public attitudes toward gay men and lesbians. *Polity, 34 (4),* 433-456.

Ownby, R.L., Czaja, S.J., Loewenstein, D. & Rubert, M. (2008). Cognitive abilities that predict success in a computer-based training program. *Gerontologist, 48(2),* 170-180.

Paivio, A. & Walsh, M. Bons, T. (1994). Concreteness effects on memory: when and why? *Journal of Experimental Psychology: Learning Memory and Cognition, 20,* 1196-204.

Park, B., Ryan, C. S., & Judd, C. M. (1992). The role of meaningful subgroups in explaining differences in perceived variability for in-groups and out-groups. *Journal of Personality and Social Psychology 63,* 553-567.

Parrott, D. J. (2008). A theoretical framework for antiGay aggression: Review of established and hypothesized effects within the context of the general aggression model. *Clinical Psychology Review, 28,* 933-951.

Parrott ,D. J., Adams, H. E., & Zeichner, A. (2002). Homophobia: Personality and attitudinal correlates. *Personality and Individual Differences, 32,* 1269–1278.

Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic bias in self-perceptions: the interplay of self-deceptive styles with basic traits and motives. *Journal of personality, 66,* 1024-1060.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598-609.

Paunonen, S. V., & LeBel, E. P. (2012). Socially desirable responding and its elusive effects on the validity of personality assessments. Journal of Personality and Social Psychology, 103, 158-175.

Petróczi, A., Aidman, E. V., & Nepusz, T. (2008). Capturing doping attitudes by self-report declarations and implicit assessment: A methodology study. *Substance Abuse Treatment, Prevention, and Methodology, 3(9).*

Pettegrew, L. S., & Wolfe, G. E. (1981). Validating measures of teacher stress. *American Educational Research Journal, 19(3),* 373-396.

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in Western Europe. *European Journal of Social Psychology, 25,* 57-75.

Plugge-Foust, C., & Strickland, G. (2000). Homophobia, irrationality, and Christian ideology: Does a relationship exist? *Journal of Sex Education & Therapy, 25.*

Rest, J. (1977). *Development in Judging Moral Issues – A Summary of Research using the Defining Issues Test.* Minneapolis, MN: University of Minnesota.

Ridgeway y, I., Roche, B., Gavin, A., & Ruiz, M. R. (2010). Establishing and eliminating IAT effects in the laboratory: Extending a behavioral model of the Implicit Association Test. *European Journal of Behavior Analysis, 11,* 133–150.

Roche, B., & Barnes, D. (1996). Arbitrarily applicable relational responding and sexual categorization: A critical test of the difference relation. *The Psychological Record, 46,* 451-475.

Roche, B., Barnes, D., & Smeets, P. M. (1997). Incongruous stimulus pairing and conditional discrimination training: Effects on relational responding. *Journal of the Experimental Analysis of Behavior, 68,* 143-160.

Roche, B., Barnes-Holmes, Y., Barnes-Holmes, D., Stewart, I., & O'Hora, D. (2002). Relational Frame Theory: A new paradigm for the analysis of social behaviour. *The Behaviour Analyst, 25(1),* 75-91.

Roche, B., O'Reilly, A., Gavin, A., Ruiz, M., & Arancibia, G. (2012). Using behavior-analytic implicit tests to assess sexual interests among normal and sex-offender populations. *Socioaffective Neuroscience and Psychology, 2,* 17335.

Roche, B., Ruiz, R. O'Riordan, M, & Hand, K. (2005). A relational frame approach to the psychological assessment of sex offenders. In M. Taylor & E. Quayle (Eds.), *New developments in child pornography and the Internet.* (pp. 109-125) Dorset: Russell House

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *The Psychological Record, 60,* 527-542.

Rohner, J. C., & Björklund, F. (2006). Do self-presentation concerns moderate the relationship between implicit and explicit homonegativity measures? *Scandinavian Journal of Psychology, 47(5),* 379-85.

Röhner, J., Schröder-Abé, M., & Schütz, A. (2011). Exaggeration is harder than understatement, but practice makes perfect! Faking success in the IAT. *Experimental Psychology, 58(6),* 464-472.

Rossion B., Joyce C. A., Cottrell G. W., Tarr M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *NeuroImage, 20,* 1609–1624.

Rothbart, M., & John, O. P. (1985). Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact. *Journal of Social Issues, 41(3),* 81-104.

Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test: Dissociating salience from associations. *Journal of Experimental Psychology: General, 133,* 139–165.

Rothermund, K., Wentura, D., & De Houwer, J. (2005). Retrieval of incidental stimulus-response associations as a source of negative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(3),* 482– 495.

Rowatt, W. C., Tsang, J., Kelly, J., LaMartina, B., McCullers, M., McKinley, A. (2006). Associations between religious personality dimensions and implicit homosexual prejudice. *Journal for the Scientific Study of Religion, 45,* 397-406.

Rowen, C. J., & Malcolm, P. J. (2002). Correlates of internalized homophobia and homosexual identity formation in a sample of Gay men. *Journal of Homosexuality, 43,* 77-92.

Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17,* 437- 465.

Rust, J. & Golombok, S. (1999). *Modern Psychometrics: The Science of Psychological Assessment* (2nd Ed.). Routledge: London and NY.

Sánchez, F. J., Greenberg, S., Liu, W. M., & Vilain, E. (2009). Reported effects of masculine ideals on Gay men. *Psychology of Men & Masculinity, 10,* 73–87.

Sandfort, T. G. M., de Graaf, R., Bijl, R. V., Schnabel, P. (2001): Same-sex sexual behavior and psychiatric disorders. *Archives of General Psychiatry, 58,* 85-91.

Schauss, S. L., Chase, P. N., & Hawkins, R. P. (1997). Environment-behavior relations, behavior therapy and the process of persuasion and attitude change. *Journal of Behavior Therapy and Experimental Psychiatry, 28,* 31-40.

Schellenberg, E.G., Hirt, J., & Sears, A. (1999). Attitudes toward homosexuals among students at a Canadian university. *Sex Roles, 40,* 139-152.

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 82-102.

Sidman, M. (1971). Reading and auditory–visual equivalences. *Journal of Speech and Hearing Research, 14,* 5-13.

Sidman, M. (1994). *Equivalence Relations and Behavior: A Research Story.* Boston: Authors Cooperative.

Simon, A. (1995). Some correlates of individuals' attitudes toward lesbians. *Journal of Homosexuality, 29(1),* 89-103.

Simoni, J. M. (1996). Pathways to prejudice: Predicting students' heterosexist attitudes with demographics, selfesteem, and contact with lesbians and Gay men. *Journal of College Student Development, 37,* 68-78.

Skinner, T. C., Blick, J., Coffin, J., Dudgeon, P., Forrest, S., & Morrison, D. L. (2013). Comparative validation of self-report measures of negative attitudes towards Aboriginal Australians and Torres Strait Islanders. *Rural and Remote Health, 13,* 1959.

Spalding, L. R., & Hardin, C. D. (1999). Unconscious unease and self-handicapping: Behavioral consequences of individual differences in implicit and explicit self-esteem. *Psychological Science, 10,* 535-539.

Stangor, C., Sullivan, L. A. & Ford, T. E. (1992). Affective and cognitive determinants of prejudice. Social Cognition, 9, 359-380.

Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology, 51,* 165-179.

Steffens, M. C., & Plewe, I. (2001). Items' cross-category associations as a confounding factor in the Implicit Association Test. *Zeitschrift Fur Experimentelle Psychologie, 48(2),* 123-134.

Steffens, M. C., & Wagner, C. (2004). Attitudes towards lesbians, Gay men, bisexual women, and bisexual men in Germany. *Journal of Sex Research, 41,* 137-149.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Tanaka, J. W., Luu, P., Weisbrod, M. & Kiefer, M. (1999). Tracking the time course of object categorization using event-related potentials. *NeuroReport, 10,* 829-835.

Taywaditep, K. J. (2001). Marginalization among the marginalized: Gay men's anti-effeminacy attitudes. *Journal of Homosexuality, 42(1),* 1-28.

Teachman, B. A., Gregg, A., & Woody, S. (2001). Implicit processing of fear-relevant stimuli among individuals with snake and spider fears. *Journal of Abnormal Psychology, 110,* 226- 235.

Tetlock, P. E., & Mitchell, G. (2009). Implicit bias and accountability systems: What must organizations do to prevent discrimination? In B.M. Staw & A. Brief (Eds.), *Research in organizational behavior* (pp. 3-38). New York: Elsevier.

Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia, 45,* 163–173.

van Well, S., Kolk, A. M., & Oei, N. Y. L. (2007). Direct and indirect assessment of gender role identification. *Sex Roles, 56,* 617-628.

Vanman, E. J., Saltz, J. L., Nathan, L. R., & Warren, J. A. (2004). Racial discrimination by low-prejudiced whites: Facial movements as implicit measures of attitudes related to behavior. *Psychological Science, 11,* 711-714.

Vuković, M. B., & Štulhofer, A. (2012). 'The whole universe is heterosexual!' Correlates of homophobia in seven southeast European countries. In K. Ringdal & Simkus, A. (Eds.). *The aftermath of war: Experiences and social attitudes in the western Balkans* (pp. 217-234). Ashgate: Abingdon.

Warriner, K., Nagoshi, C. T., & Nagoshi, J.L. (2013). Correlates of homophobia, transphobia, and internalized homophobia in Gay or lesbian and heterosexual samples. *Journal of Homosexuality, 60(9),* 1297-1314.

Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41,* 33-50.

Weinberg, G. (1972). *Society and the Healthy Homosexual.* Garden City: New York Anchor.

Whatmough, C., Verret, L., Fung, D., & Chertkow, H. (2004). Common and contrasting areas of activation for abstract and concrete concepts: an $H_2{}^{15}O$ PET. *Journal of Cognitive Neuroscience, 16,* 1211-1226.

Wigboldus, D. H. J., Holland, R. W., & van Knippenberg, A. (2005). Single Target Implicit Associations. Unpublished manuscript.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120,* 3–24.

Wilson, T. D. (2002). *Strangers to Ourselves: Discovering the Adaptive Unconscious.* Cambridge, MA: Belknap Press/Harvard University Press.

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

Yang, A. S. (1997). The polls-trends: Attitudes towards homosexuality. *Public Opinion Quarterly, 61,* 477-507.

Ziegert, J. C., & Hanges, P. J. (2005). Employment discrimination: The role of implicit attitudes, motivation, and a climate for racial bias. *Journal of Applied Psychology, 90,* 553-562.

Appendix 1

Consent Form

In agreeing to participate in this research I understand the following:

This research is being conducted by Aoife Cartwright, a postgraduate student at the Department of Psychology, National University of Ireland Maynooth. It is the responsibility of this student to adhere to professional ethical guidelines in their dealings with participants and the collection and handling of data. If I have any concerns about participation I understand that I may refuse to participate or withdraw at any stage.

I have been informed as to the general nature of the study. I understand that as a requirement of participating in the study I will be exposed to a computer-based task which will involve the presentation of images/words, some of which are sexual in nature. I am also happy to complete a series of questionnaires that will ask me questions about my attitude to homosexuality and sexuality in general.

All data from the study will be treated confidentially and my data will not be identified by name at any stage of the data analysis or in the final report. The data will be compiled, analysed and submitted in a report to the Psychology Department, NUI, Maynooth.

I understand that no clinical judgement can be made of me on the basis of my participation or performance during this research and that because this is a group-based study my own individual responses on the questionnaire and the computer based tasks are of no interest to the researchers.

At the conclusion of my participation, any questions or concerns I have will be fully addressed.

I may withdraw from this study at any time, and may withdraw my data at the conclusion of my participation if I still have concerns.

I am over 18 years of age.

Signed:

_____Participant

_____ Researcher

_____ Date

Appendix 2

MHS

Please circle the number that best represents your level of agreement with each of the following statements.

Participant Number _____

Participant Gender _____

Participant Age _____

1. Many Gay men use their sexual orientation so that they can obtain special privileges.

Strongly disagree 1 2 3 4 5 Strongly agree

2. Gay men seem to focus on the ways in which they differ from heterosexuals and ignore the ways in which they are the same.

Strongly disagree 1 2 3 4 5 Strongly agree

3. Gay men do NOT have all the rights they need.

Strongly disagree 1 2 3 4 5 Strongly agree

3. The notion of universities providing students with undergraduate degrees in Gay and Lesbian studies is ridiculous.

Strongly disagree 1 2 3 4 5 Strongly agree

4. Celebrations such as “Gay pride day” are ridiculous because they assume an individual’s sexual orientation should constitute a source of pride.

Strongly disagree 1 2 3 4 5 Strongly agree

5. Gay men should stop shoving their lifestyle down other people’s throats.

Strongly disagree 1 2 3 4 5 Strongly agree

6. Gay men should stop complaining about the way they are treated in society and simply get on with their lives.

Strongly disagree 1 2 3 4 5 Strongly agree

7. Gay men have become far too confrontational in their demand for equal rights.

Strongly disagree 1 2 3 4 5 Strongly agree

8. In today's tough economic times, tax payers' money should not be used to support Gay organizations.

Strongly disagree 1 2 3 4 5 Strongly agree

9. If Gay men want to be treated like everyone else then they need to stop making such a fuss about their sexuality or culture.

Strongly disagree 1 2 3 4 5 Strongly agree