Developing the Function Acquisition Speed Test: Using a Functional Research Approach to Build a Novel Implicit Test

Department of Psychology

National University of Ireland, Maynooth

PhD Thesis

Thesis submitted for Doctoral Degree

Submitted by Anthony O’Reilly

Supervised by Dr. Bryan Roche

February 2012

Table of Contents

Acknowledgements

Today, I sit down to write this and realise that I am a very lucky man. Countless seemingly small decisions and events brought me down this road, into contact with the many fantastic people who have helped me to this point. At this milestone in my life, ten years in the making, I want to try to articulate just how lucky I feel to have had you all in my life. I can't thank you all as much as you deserve, but I'll give it a go anyway.

To Bryan - Six and a half years ago, I sat down with you to talk about a final year project. A few short weeks later, I had been bitten by the research bug, and there was no going back. It may have taken me three years before I returned to NUI Maynooth, but looking back, it was inevitable. Your infectious enthusiasm for learning, for research, and for a good argument/rant/tangent brought out the best of the same traits in me, and I couldn't imagine tackling a PhD with anyone else as my supervisor. No matter what mood I was in when I walked through your door, I'd come out bursting with energy and enthusiasm. I hope I can pass the same on to my students in the future, because I cannot articulate how valuable that was to me.

To all of the staff at NUI Maynooth – Coming to study Psychology here in 2003 was one of those "small" decisions. It may surprise you to learn that teenaged Anthony was not a paragon of forward thinking or planning (I even came within a hair's breadth of choosing to study English at Trinity.) I couldn't have hoped to find better teachers. Thank you, one and all.

To Louise – Every moment spent with you has been an oasis of calm and joy without compare. Thank you for every wonderful day, and every one to come.

To my friends – I am lucky to know every one of you, and each of you has in some way or another kept me sane through this. Thank you.

To my parents – As you well know, I'm not a man of few words, but I don't have half enough to thank you as much, or as well, as you deserve. You gave me everything I needed to get to where I am today, and you gave it without question or reservation. You were my first teachers, and my best.

Publications arising from research described in the current thesis

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*. doi:10.3402/snp.v2i0.17335

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 (3)*

O'Reilly, A., Roche. B., Ryan. A., & Campion, G. (In press). A Function Acquisition Speed Test for Equivalence Relations. *The Psychological Record.*

Abstract

The eleven studies reported in this thesis outline the development of a novel implicit test for assessing verbal and social histories. This measure was named the "Function Acquisition Speed Test" (FAST). The current research utilizes a functional research approach drawing upon the seminal research by Watt, Keenan, Barnes and Cairns (1991) and upon more recent research by Gavin, Roche and Ruiz (2008) to inform the bottom-up development of the FAST. Chapter 1 presents a review the behaviour analytic literature concerned with the phenomena of stimulus equivalence and derived relational responding, and reviews the links between these concepts and complex human behaviour such as attitudes. The seminal study by Watt et al. (1991) is also described. In that study, the researchers attempted to train subjects to form two equivalence classes which were incongruent with Northern Irish subjects' socioverbal history (i.e. containing Catholic names and Protestant symbols) using a stimulus equivalence paradigm. Watt et al. concluded that pre-existing social histories can interfere with the acquisition of novel stimulus relations. The relevance of this work as an underpinning principle of the FAST is described. Chapter 1 also briefly reviews the social cognitive approach to implicit testing and discusses the value of a functional approach.

Also in Chapter 1, the initial structure of the FAST paradigm is outlined. The key FAST test blocks present subjects with one stimulus per trial, and require subjects to learn to emit one of two responses (i.e. press right or press left) based on corrective feedback. In the consistent block, the same response is required for the two stimuli suspected to be related, while the other response is required for two novel, unrelated stimuli. In the inconsistent block, the two stimuli of interest require two different responses. The reinforcement contingencies of the inconsistent act as a behavioural disrupter, and the difference in learning rates between the two blocks is taken to be indicative of the strength of the relation between the test stimuli.

Chapter 2 reports on five experiments that tested the underlying premise of the FAST with regard to experimentally trained stimulus relations. In each of the experiments in this chapter, subjects completed training procedures to establish a stimulus relation which would later be probed for using the FAST methodology. In Experiment 1, subjects completed matching to sample training which established a

simple A-B relation between two three-letter nonsense syllables. The FAST was capable of detecting the trained relations in a majority of subjects and on a group level, thus establishing a basic proof of concept for the methodology. Experiment 2 sought to expand on this by establishing aversive or erotic stimulus functions for a pair of nonsense syllables. The FAST procedure was then used to test for a relation between a nonsense syllable and related (but novel) romantic images. The FAST was also run multiple times for each subject. The results revealed a great deal of instability over time, with the expected effect only emerging on the second of three FAST runs.

In Experiment 3, subjects were trained using a Matching-to-Sample procedure to create two three-member equivalence classes. The derived relations (A1-C1) were then tested for using the FAST. Given the importance of derived relational responding to a behavioural account of attitudes, this was a vital test for the fledgling format. The FAST was successful in detecting the derived relations, and this effect emerged on all three runs of the FAST. Experiment 4 sought to further build on this success by utilising an extremely robust and extended training procedure to maximise the probability of the emergence of stable equivalence relations. The training procedure contained nine stages of increasing difficulty and fading of consequences. A repeated single subjects design was used, and 4 out of five subjects showed strong positive FAST effects when tested.

Experiment 5 investigated the question of whether the FAST was capable of detecting the influence of derived relations which had not yet been tested for by a normal equivalence testing procedure. All subjects completed an equivalence training procedure, but half of the experimental subjects completed a FAST prior to completing an equivalence test, while the other half completed a FAST after completing a normal equivalence test. The FAST was unable to detect the derived relations in the former experimental group, even in subjects who latter passed the equivalence test. The implications of this result for a Relational Frame Theory account of implicit attitudes are discussed.

In Chapter 3, the effect of a number of basic procedural modifications to the FAST presentation was considered. In each experiment, subjects completed 3 successive runs of a FAST after having been trained in a simple A-B stimulus relation. Experiment 6 used a between subjects design to measure the effect of varied response windows on FAST performance (1000ms, 2000ms, 3000ms or 4000ms). Experiment 7 investigated the effect of including a variable number of practice blocks prior to the first baseline block (3, 5, or 7 practice blocks). Experiment 8 varied the fluency criterion for completing a FAST block. (7, 8, or 9 correct in a sequence of ten responses.). The results of these experiments were difficult to interpret due to instability in the predicted FAST effect across the three runs of the FAST. Possible reasons for this instability are discussed, and improvements to the FAST are suggested. The results of these experiments resulted in a shortening of the FAST response window to 2000ms, the inclusion of a practice block, and the inclusion of a "counter" which displayed the number of correct responses in a row that the subject had emitted.

Chapter 4 tested the FAST with "real world" stimuli as the test stimuli. In each experiment in Chapter 4, subjects complete FAST procedures aimed at detecting natural histories of verbal behaviour. In each experiment, a different relation is targeted. In experiment 9, the FAST probes for a relation between the words "spider" and "disgust". Experiment 10 targets the relations "immigrant" and "cheat". Experiment 11 departs from using words as stimuli, using images of teenaged females to probe for relations between sexual images and images of teenaged females of different ages. The results of these experiments demonstrate both the limitations and strengths of the FAST procedure in different contexts. The FAST was shown to be most capable of detecting culturally ubiquitous verbal relations (spiders are disgusting); particularly in cases where the two stimuli are strongly in opposition (i.e. young girls are sexual). Experiment 11 also demonstrated the potential of the FAST to be deployed in serial to examine the "shape" of an equivalence class. Subjects in experiment 11 showed a strong *negative* FAST effect when female stimuli were pre-teenaged, but this effect quickly vanished as the ages of the pictured females increased into the teenage years.

Chapter 5 provides an overview of the entire research project. The experimental findings from each of the previous chapters are discussed along with the implications of these findings. The early development of this novel test format is charted, and remaining challenges that must be confronted in future research are outlined. Alternative experimental preparations emerging from this research are also considered and explored, and possible applications in both basic research and applied setting are discussed.

Chapter 1

An Introduction

In the past two decades, the field of Behaviour Analysis has begun to rise to the challenge of answering questions about language and cognition that were thought by many to be outside the reach of its methods and research philosophy. The first behavioural account of language was offered in 1957 by B.F. Skinner in *Verbal Behaviour.* However, this account is widely considered to be inadequate, and ultimately failed to provide a fruitful program of empirical research (Dymond & Alonso- Alvarez, 2010). However, over the following two decades, further research into basic behavioural processes yielded fresh understandings of the differences between the behaviour of humans and non-humans which have informed a new theoretical approach to language and cognition from a behavioural perspective.

In particular the *Relational Frame Theory* (RFT: Hayes, Barnes-Holmes & Roche, 2001) approach has led to a multitude of advances in behavioural understanding of complex human language and cognition. Areas which have benefited from the RFT framework include analogical reasoning (Carpenter, Smeets & Barnes-Holmes, 2003; Stewart & Barnes-Holmes 2004; Stewart, Barnes-Holmes & Weil, 2009), assessment and training of intelligent performance (e.g., O'Toole & Barnes-Holmes, 2009; Cassidy, Roche & Hayes, 2011), perspective taking in applied developmental and clinical arenas (McHugh, Barnes-Holmes & Barnes-Holmes, 2004; Rehfeldt, Dillen, Ziomek & Kowalchuck, 2007; Villatte, Monestes, McHugh, Freixa i Baque & Loas, 2010a & b; Weil, Hayes, & Capurro, 2011) and generative verbal behaviour in developmentally delayed children (e.g., Moran, Stewart, McElwee & Ming, 2010; Murphy & Barnes-Holmes 2009; Heagle & Rehfeldt, 2006). The focus of the current thesis is the assessment of attitudes, another area in which these new advances in this field can be of great utility.

Relational Frame Theory offers the attitude researcher an account of language (and thus cognition) that is founded upon a single core process – relational framing. This process is precisely articulated in a bottom up account which specifies the interactions between organism and environment in terms of the history of a single organism. The bottom-up, functional analytic approach taken by RFT builds incrementally from elementary core processes, but retains the explanatory power necessary to tackle more complex forms of behaviour with the same precise terminology. The relational framing account of language has the potential to shed light on the fundamental processes which underlie attitude constructs, allowing for a more nuanced theoretical account of their formation and change, and to guide research into methodologies which might more accurately and reliably measure attitudes.

From Stimulus Equivalence to Arbitrarily Applied Relational Responding

Much of the research that underpins current thinking on relational framing has grown from the discovery of the phenomenon of stimulus equivalence (Sidman 1971). Sidman's investigations began while researching participants who experienced difficulty reading, writing, and speaking. As such, it was necessary for Sidman and his colleagues to develop procedures that could investigate their language behaviour without requiring them to speak or write. To achieve this, Sidman to the Matching-to-Sample (MTS) procedure, a conditional discrimination procedure first developed to study nonhuman behaviour (e.g. Yerkes, 1928). In an MTS procedure a single *conditional stimulus* (either A1 or A2, varied between trials) is presented to the participant. Then, two further stimuli (B1 and B2) are presented as response options to be choose between. The correct response is determined by the which of the two conditional stimuli are presented, thus leading to an "if-then" relation such as "If A1 is present, select B1. If A2 is present, select B2". In Sidman's version of the MTS task, each trial involved the presentation of a sample stimulus, either a picture of the object to be named (e.g. a picture of a cat), a word (e.g. cat"), or an auditory stimulus (e.g. the word "cat" spoken aloud to the participant). In matching tests, the participants were required to press on the correct comparison stimulus (a picture or a word) from an array of eight choices (pictures or words). In oral naming tests, the participant was required to name the sample stimulus aloud.

The participant came to the experiment being able to choose the correct picture when a word was spoken to him, and to name an object aloud when the picture was presented – he had good auditory comprehension but poor reading comprehension. Using the matching to sample procedure, the researchers taught the participant to match the spoken word samples to the correct written word. Without any further direct training, the participant was then able to match written words to pictures (and vice versa) and to name the written words. This emergent behaviour caused significant excitement in Sidman and his collaborators (Sidman, 1982). The vital finding was that teaching two sets of conditional discriminations caused novel behaviours to emerge without direct training. This early glimpse of the phenomenon caused Sidman to focus his research on defining stimulus equivalence and establishing the necessary and sufficient conditions required to produce and test it in the laboratory.

Stimulus equivalence was defined procedurally (Sidman, 1982) as responding which displays the properties of reflexivity, transitivity, and symmetry. When a verbally able human participant is trained in (at least) two conditional discriminations (e.g. given A1 pick B1 and not B2, given A2 pick B2 and not B1, given B1 pick C1 and not C2, and given B2 pick C2 and not C1) the participant will behave in ways that have not been reinforced by the experimenter. When a participant is presented with A1, he will pick A1, matching each stimulus with itself (reflexivity.) When a participant is presented with B1, he will pick A1, reversing the direction of the trained relation (symmetry.) When presented with C1, he will select A1, deriving the untrained identity relation between the stimuli that were never paired (transitivity).

The generativity and stimulus substitutability characteristic of Stimulus Equivalence suggested a strong link between stimulus equivalence and language. Verbal behaviour is thought to be the source of most of the differences between humans and non-humans with regard to their responses to different learning preparations (Languages Hypothesis; Lowe, 1979.) Early attempts to explain verbal behaviour (Skinner, 1974) relied entirely on operant conditioning (in the absence of stimulus equivalence) as the process for learning verbal behaviour, and commentators did not believe that this was adequate to explain the range and complexity of human verbal behaviour. However, with the emergence of research into stimulus equivalence, behaviour analysts had a new working model of verbal relations (Sidman & Tailby, 1982). The evidence that links stimulus equivalence behaviour to language is compelling.

Nonhumans do not readily demonstrate stimulus equivalence in the laboratory. Indeed, it is debatable as to whether they show stimulus equivalence at all (Hayes, 1989) despite attempts by researchers to do so with birds (D'Amato, Salmon, Loukas, & Tomie, 1985; Lipkens, Kop, & Matthijs, 1988, Vaughn, 1988) and primates (D'Amato et al.,1985 Macintyre, Cleary & Thompson, 1987; Sidman et al., 1982). Whether this failure is due to methodological problems or due to a uniquely human capacity for stimulus equivalence is uncertain. What is certain is that stimulus equivalence does not occur readily in nonhumans. Performance on relational and stimulus equivalence tasks appears to correlate with verbal ability (Barnes, McCullagh & Keenan, 1990; Devany, Hayes & Nelson, 1986; Dugdale and Lowe, 1990; O’Hora, Pelaez, and Barnes-Holmes 2005; Cassidy et al., 2011.)

Research has also been conducted into the neurophysiological correlates of derived relation responding and stimulus equivalence. Barnes-Holmes and colleagues (Barnes-Holmes et al., 2005) examined Event Related Potentials (ERPs) associated with analogical reason from an RFT perspective. Their experiment showed electrophysiological activity in the dorsolateral prefrontal areas associated with both similar-similar and different-different relational responding. RFT would predict that the former is "simpler" and functionally distinct from the latter – in the case of similar-similar responding, only one relational frame (equivalence) is involved. In line with predictions, waveforms were significantly more negative in the left-hemispheric pre frontal regions for different-different relating. These findings were consistent with neurocognitive literature on analogy, and as such strengthen the RFT model with regard to analogical reasoning.

fMRI research from a variety of researchers has also shed light on what may be the fundamental neurocircuitry underlying Derived Relational Responding (Dickens et al., 2001, Schlund et al., 2008, Hinton et al., 2010). The core set of brain regions implicated are the prefrontal cortex and the posterior parietal regions. Cognitive neuropsychology research has identified these brain regions as being involved in the regulation of cognitive and behavioural responses. It has been further suggested that these brain regions represent a uniquely human fronto-parietal network (Ogawa et al., 2009) whose architecture facilitates equivalence responding in humans as opposed to animals.

Stimulus equivalence also offers an explanation for the generativity of language. Stimulus equivalence demonstrates a process by which from two trained relations (A-B and B-C) four more untrained relations emerge (A-C, C-A B-A, C-B). Equivalence responding can emerge across classes far larger than 3 members, and in these cases, the number of derived relations generated increases rapidly. In a 4 member equivalence class, 3 trained relations (A-B, B-C, C-D) generate nine derived relations (A-C, A-D, B-D, B-A, C-B, D-C, C-A, D-A, D-B). With a five member equivalence class, four trained relations generate sixteen derived relations (i.e., the number of derived relations is equal to the square of the number of trained relations). When this is considered in light of the number of relations trained by the verbal community in even a short time, the amount of emergent complexity ought to be self evident. It is worth noting that the emergence of equivalence behaviour at around age two (Devany et al., 1986) coincides with the so called "language explosion" when a child's repertoire of learned words increases greatly. Given this generativity and emergent complexity, it becomes possible to see how an account based on simple core principles could still retain the explanatory power to tackle highly complex and changeable language constructs such as attitudes.

The Relational Frame Theory Account

With the preceding in mind, it should come as no surprise that behaviour analysts have pursued stimulus equivalence research with vigour and have embraced it as potentially being the core process in language behaviour. However, the procedural definition of equivalence offered by Sidman is not a definition of the behavioural process itself (Sidman, 1992; Barnes, 1994.) Rather, the particular experimental preparation brings the process under rigorous experimental control and allows the experimenter to be quite certain that equivalence is occurring. What was needed, then, was a theoretical account of stimulus equivalence that could guide the burgeoning behavioural research into complex human behaviour and language.

A Relational Frame Theory (RFT; Hayes, Barnes-Holmes & Roche, 2001) account of language and cognition draws upon and elaborates the stimulus equivalence phenomenon described above.

As well as being able to discriminate (i.e., detect and respond to) specific stimuli, organisms are also capable of responding to relations *between* stimuli, such as similarity, difference, distance, greater than, and so on. This is known as *relational responding*. Nonverbal organisms are capable of learning to respond to such *formal* relations, such as size and distance, via traditional learning processes such as operant conditioning (see Reese, 1968). Verbal organisms, however, display the unique ability to respond to *arbitrary* stimulus relations such as oppositeness, value and time. This form of responding is called *arbitrarily applied relational responding.* (AARR; Hayes, Barnes-Holmes & Roche, 2001)

As seen in stimulus equivalence, verbal organisms can derive equivalence between stimuli that has never been explicitly trained. The different forms of *relational responding* (difference, opposition, greater than, less than, etc.) can also be derived without explicit training (Dymond and Barnes, 1995; Lipkens, Hayes, & Hayes, 1993; Roche & Barnes, 1996). Derivations of relations other than equivalence require specific nomenclature to allow researchers to speak with precision about the relations under analysis. Rather than “stimulus equivalence” RFT researchers use the phrase “derived relational responding” (DRR) to refer to both the stimulus equivalence phenomenon, and the wider ranged of possible relations that can be learned and derived. Different nomenclature is also used to describe the properties of DRR. *Mutual entailment* means that if A is related to B, then B is related to A in a complimentary fashion. For example, if A is opposite to B, then B is opposite to A. If A is more than B, then B is less than A. *Combinatorial entailment* occurs when three or more stimuli are related. If A is opposite to B, and B is opposite to C, then the relation that is derived between A and C is one of equivalence, because both are opposite to B. Combinatorial entailment refers to the reciprocal relationships that exist between two stimuli as mediated by other intermediary stimuli (Blackledge, 2003).

A vital feature of the RFT account with regard to the discussion of attitudes is the finding that the psychological functions of any stimulus will be transformed by its relation to other stimuli in a derived relational network (*Transformation of function*). The concept of Transfer of Function has led to the emergence of the novel behaviour analytic formulation of attitudes described below (Grey & Barnes, 1996; see also Moxon, Keenan, & Hine, 1993;

Roche, Barnes-Holmes, Barnes-Holmes, Stewart, & O'Hora, 2002; Schauss, Chase & Hawkins, 1997).

An attitude may be conceived of as a network of derived and explicitly reinforced stimulus relations according to which the functions of events are transformed (e.g., Grey and Barnes, 1996). As such, an attitude can be constructed as a verbal event (or series thereof) which emerges from our interactions with others and with our environment. From this perspective, an attitude is established and maintained through a history of both explicitly reinforced relations and untrained derived relations between verbal stimuli. The remote social contingencies which support these networks of relational responses can be represented by the verbal practices of the wider community (i.e. the culture with which the participant is in contact.) For instance, rules, norms, mores, and taboos all constitute forms of verbal contingency that specify relations between stimuli (e.g., "sex" and "dirty") in often complex and subtle (i.e., indirect) ways. A participant's past participation in a verbal environment (i.e., a culture) provides many hundreds of training exemplars that establish complex derived relational networks through which functions of stimuli may be transformed, and this can explain the emergence of apparently untrained or indirectly trained responses and attitudes. This occurs because of the way in which the various terms were framed relationally in language, whether explicitly, or in turn by further derived relations (e.g., innuendo, jokes). This approach has a great deal in common with social constructionist and feminist theories regarding the social construction of gender roles and similar socially prescribed roles and norms (Roche et al., 2002).

Kohlenberg, Hayes, and Hayes (1991) investigated a model of social stereotyping based on verbal control over equivalence classes. The example provided was as follows; in the sentence "the woman complained and complained", the word "woman" may serve as a contextual cue for relating the word "complain" to the words "*nag"* or "*bitch".* In a structurally similar sentence “the man complained and complained”, the word “man” serves as a contextual cue which occasions relating “complain” with “*assertive”* or “*forceful”*. In their Experiment 1, participants were exposed to conditional discrimination to form six four-member equivalence classes. The contextual stimuli consisted of six gender-identifying names (three male, three female). After testing for equivalence using the original six names, the participants were then tested using novel male and female names. Control of the conditional discriminations successfully generalised via equivalence to the novel male and female names. Experiment 2 brought the process under further control by generating pre-experimental equivalence classes (rather than simply using male/female names). Stimuli from these classes were then used as contextual stimuli in a conditional discrimination task of the same type as used in Experiment 1. In line with predictions, contextual control transferred to the rest of the pre-trained equivalence classes, just as they had with the male/female names in Experiment 1. These experiments showed that contextual control could transfer without direct training from a stimulus to other stimuli with which it shares an equivalence class. These experiments modelled the stereotyping that underpins the concept of attitudes.

Grey and Barnes (1996) suggested that a negative attitude towards normal heterosexual interactions can be seen as responding in accordance with an equivalence relation between normal opposite-sex adults and descriptive terms such as 'disgusting'. Their empirical study demonstrated a transformation of a trained attitudinal or evaluative response from one member of an equivalence class to the other members of the equivalence class.

Grey & Barnes (1996) provided participants with the necessary conditional discrimination training to form the following three derived equivalence relations; A1-B1-C1, A2-B2-C2, and A3-B3-C3, using nonsense syllables as stimuli. Participants then viewed the video contents of VHS video cassettes that were clearly labelled with the A1 and A2 stimuli. One of the cassettes contained sexual/romantic scenes, while the other contained religiously themed scenes. Subsequently, participants were asked to categorize four novel video cassettes, each labelled as B1, C1, B2 or C2. They were given no information about these cassettes, but they categorised them according to the derived equivalence classes. That is, participants classified the B1 and C1 cassettes in the same way as the A1 cassette, and the B2 and C2 cassette in the same way as the A2 cassette. In effect, the study demonstrated the sexual and religious evaluative functions of the A-labelled cassettes and the derived relations in which the A stimuli participated, transformed the functions of the C-labelled cassettes, such that these were responded to as sexual or religious as appropriate. This study demonstrated a process by which evaluative responses (i.e. attitudes) can be indirectly trained, providing support for the behaviour analytic model of attitudes described above.

In a similar study, Barnes and Roche (1997) attempted to generate a derived laboratory induced fetish to extend the work of Rachman (1966). Participants were trained to emit a sexual arousal response to a stimulus, which transferred to other members of an equivalence class in the absence of explicit training. In a theoretical paper from the same authors (Barnes & Roche, 1997) it was suggested that such a transfer of functions may account for behaviours in the real world, particularly those of a sexual nature (see also Roche & Barnes, 1998).

Dixon, Dymond, Rehfeldt, Roche, & Zlomke's (2003) applied a transfer-of-function model to the understanding of the September 11th terrorist attacks in the USA. On hearing of the terrorist attack (A) our white American male instantly experiences feelings of rage (B). The media claim that Terrorists are responsible for these acts, and depicts pictures of these Terrorists on the television (C). The images of the terrorists themselves may now come to elicit feelings of hate or rage through a transfer of function across the stimuli in the newly created relation. As the most salient features of the terrorists are their race, religion, and country of origin these feelings of hate and rage towards the terrorists begin to transfer to other persons sharing the same skin colour, religion, and country of origin because of a formal similarity between them and the terrorists. That is, innocent Muslims of a Middle Eastern descent are now added to the A-B-C relation as a fourth stimulus “D”. The formal properties of the B stimuli and the D stimuli are approximately the same through transfer of stimulus functions. Moreover, the feelings of hate and rage held by our American white male were initially occasioned only by the terrorist attacks now have transferred beyond the terrorists themselves.

Detecting Histories of Verbal Behaviour: The Watt et al. Paradigm

The pivotal role played by derived relational responding in language led researchers to investigate how histories of verbal behaviour might impact upon participants’ formation of new equivalence classes. In their seminal study, Watt, Keenan, Barnes and Cairns (1991) used a simple stimulus equivalence paradigm in which participants were trained to relate stimuli with strong socially established functions in ways which were inconsistent with their social history. Specifically, they took advantage of the fact that people in Northern Ireland often respond to names as indicative of religious background, and utilised stimuli representative of Catholic and Protestant names and symbols.

A three-phase experimental procedure was employed. Participants were first exposed to a matching to sample procedure comprising the presentation of either a nonsense syllable or a first and last name at the top of the screen (the "sample" stimulus). Three "comparison" stimuli were displayed separately below. Participants were instructed to select a comparison stimulus by pressing a corresponding key. Training comprised of one of three Catholic names being randomly chosen to serve as the sample stimulus. Beneath this, three nonsense syllables served as comparison stimuli and were arranged in a random order across the screen. Participants were required to select the correct comparison in the presence of the sample stimulus (A-B Relations.) The second stage trained B-C relations. Here, the sample stimuli were selected from the list of nonsense syllables, and the comparison stimuli were selected from the list of Protestant symbols. Feedback was provided on all trials during Stage 1. During stage 2, corrective feedback was presented on 50% of responses. The stimulus combinations described in Stage 1 were all presented in random order during this condition. Each stimulus combination was presented twice and participants were required to meet 100% criterion.

Stage 3 of the Watt et al. procedure involved Equivalence Testing. For this stage, no corrective feedback was provided. Ten presentations of each of the stimulus combinations from Stage 1 were randomly presented. Interspersed with these were ten presentations each of six other stimulus combinations. Each of the three Protestant symbols served as sample stimuli and two of the Catholic names served as comparison stimuli. An additional Protestant name was included as a comparison stimulus for each of these three combinations of sample and comparison stimuli. The results of the Watt et al. study showed that during equivalence testing, all of the English participants correctly matched the Catholic names with the Protestant symbols, but 12 of the 19 Northern Irish participants chose a novel Protestant name in the presence of the Protestant symbols, thereby failing to respond equivalently. These findings strongly suggested that the social contingencies operating in Northern Ireland interfered with the establishment of equivalence relations in the laboratory. The equivalence test required Northern Irish participants to juxtapose names and symbols in a manner that was counter-cultural for this group of participants. As such, the Watt et al. procedure showed potential for a stimulus equivalence paradigm to underpin a test which could assess social history without the participant being able to identify the nature of the task.

Similar studies supported the suggestion made in the Watt et al. study that social history interferes with the formation of equivalence classes. In a study on gender identity

Moxon et al. (1993) found that participants had more difficulty forming equivalence classes when the classes included female names and stereotypic male occupations. Leslie, Tierney, Robinson, Keenan, Watt, and Barnes (1993) also employed the Watt et al. procedure in a study with which examined the paradigm could be used to differentiate between anxious and non-anxious participants. Results demonstrated that participants in the anxious group more difficulty in matching pleasant-state adjectives to threatening situations. Merwin and Wilson (1995) required participants to form equivalence classes between self referential terms and negative evaluations, and in a second procedure, between those terms and positive evaluations. Participants who reported high distress and low esteem hade significantly more errors when required to match “self” terms with positive items. Plaud (1995) showed that participants required significantly more training to establish equivalence classes made up of aversive stimuli (snake-related words) than classes made up of innocuous stimuli (flower related words). Importantly, the increase in the amount of training required to establish the equivalence classes correlated with self reported fear of snakes.

These studies supported the assertion that pre-experimental functions of stimuli could interfere with the formation of equivalence classes in the laboratory. However, in order to explore the processes involved fully it was necessary for researchers to conduct research using experimentally established functional classes and examine their effect on the formation of equivalence relations.

The first study to address this issue was Roche, Barnes, and Smeets (1997). The experimenters trained participants on a matching to sample procedure that formed two three-member equivalence classes using nonsense syllables as stimuli (i.e. A1-B1-C1 and A2-B2-C2.) The authors then paired two of the stimuli (A1 and C2) with sexually arousing film clips and two other stimuli (A2 and C1) with nonsexual film clips, establishing conflicting sexual arousal functions for stimuli that had been trained as equivalent. When re-exposed to the equivalence testing, participants reproduced the original equivalence responses. In a second experiment, Roche et al. exposed established the sexual/nonsexual functions of A1, C2, A2 and C2 first, and then presented participants with the matching-to-sample procedure. Participants matched stimuli based on their conditioned sexual/non-sexual functions (i.e. A1 with C2 and A2 with C1) rather than forming the equivalence classes required by the matching to sample procedure. This demonstrated that once an equivalence class is formed, it is difficult to disrupt with succeeding functional relations, and that the reverse is also true – it is difficult for participants to form equivalence relations when they are incongruous with existing functional relations.

The issue was further examined by Tyndall, Roche and James (2004). They established two functional classes of stimuli – Six S+ stimuli (responding to the stimulus was reinforced) and Six S- stimuli (responding away from the stimuli was reinforced.) In matching to sample training, participants were trained to form two three-member equivalence classes using four different combinations of S+/S- stimuli. Participants required more training to establish two distinct equivalence classes from amongst 6 S+ stimuli (i.e. functionally similar stimuli) than from amongst 6 S- stimuli (i.e. without a shared function.) Further, participants found it easier to form equivalence classes when they were required to separate S+ and S- stimuli than when they were required to form classes which mixed S+ and S- stimuli.

Tyndall et al. (2009) extended the previous work. While in the 2004 paper, the stimulus functions were emotionally neutral, the later study established functional classes by pairing six stimuli with aversive images and six stimuli with neutral images, and then tested for the formation of those functional classes. In equivalence training and testing, the authors found that participants required significantly more training to establish two three-member equivalence classes from amongst the six aversive stimuli than from amongst the six neutral stimuli.

Given the preceding research, it should come as no surprise that behaviour analysts began to explore the potential of using the Watt et al. paradigm as a basis for an experimental measure that would be sensitive to pre-experimental verbal histories. McGlinchey, Keenan and Dillenburger (2000) examined the extent to which normal equivalence responding can be disrupted by socially loaded stimuli. A group of children first participated in a standard equivalence training and testing procedure, using nonsense syllables and a range of pictures. Each child subsequently took part in a dressing-up role play in which the photographed hat, goggles, braces and shirt were employed. In order to socially load the clothing items and related stimuli, some of the clothes were purposely placed on inappropriate body parts. This was intended to recombine the relations indirectly between the stimuli in the naturalistic manner in which a child might acquire confusing or inappropriate information during an abusive episode. Each child was then re-exposed to the equivalence test. It was expected that equivalence responding (e.g., matching B1 to B2) would be disrupted following the role play. The final equivalence test revealed patterns of responding that were sensitive to the inappropriate information that the children had knowingly or unknowingly acquired during the dress-up role play.

The Social-Cognitive Approach

At the same time that behaviour analysts were investigating the potential of the Watt et al. procedure for creating a test to detect a participant’s history of verbal behaviour, a new research thread was opening up in social cognitive psychology which bore several similarities to the work already described.

Drawing upon research in implicit memory, Greenwald and Banaji (1995) outlined the idea of *implicit social cognition.* The authors described how past experience can influence present attitudes and stereotypes and the responses mediated by these constructs. Core to the concept was the idea that some experiences, while they influence behaviour, are not available to accurate introspection or self report. (Note the similarity between this idea and the performance of participants in the Watt et al. procedure). They set out to provide a measurement method that would be able to detect these hypothetical implicit constructs accurately, which they dubbed the "Implicit Association Test." For the social cognitive psychologist, indirect "implicit" measures" offered a new and exciting tool for research, and they were embraced with a great deal of enthusiasm.

Several implicit tests have been created to date; The Evaluative Priming Task (Fazio, Jackson, Duntan & Willians, 1995), the Semantic Priming Task (Wittenbrink, Judd & Park, 1997), the Affect Misattribution Task (Payne, Cheng, Govorun, & Stewart, 2005), the Extrinsic Affective Simon Task (De Houwer, 2003a), and the Go/No Go Association Task (Nosek & Banaji, 2001). However, the IAT has been overwhelmingly the most widely utilized. The IAT has been utilised across many different fields of study within psychology, including (but not restricted to) forensic psychology (Brown, Grey and Snowdon, 2009), clinical psychology (Egloff and Schmulke, 2002) and health psychology (e.g., Teachman, Gapinski, Brownell, Rawlins, & Jeyaram, 2003; Von Hippel, Brener, & von Hippel, 2008).

The underlying assumption of the IAT is that it is easier to assign a single response to two concepts if they are associated in memory than if they are unrelated. Greenwald's initial experiment presented participants with flower names (e.g. TULIP) insect names (e.g. SPIDER), pleasant words (e.g. LOVE) or unpleasant words (e.g. UGLY) and asked to categorize them by means of a key press.

One early IAT study (Greenwald, McGhee, & Schwartz 1998) involved presenting participants with flower names (e.g. TULIP) insect names (e.g. SPIDER), pleasant words (e.g. LOVE) or unpleasant words (e.g. UGLY) individually on separate trials. Participants were asked to categorize them by means of one of two positional key-presses. In the first (consistent) condition, the same response key was assigned to both flower and pleasant stimulus words, while insects and unpleasant word stimuli shared a different but common positional response. In the second (inconsistent) condition, one response key was assigned to unpleasant words and flower stimuli, and the other to pleasant words and insect stimuli. The researchers found that reaction times were shorter for associated stimuli (i.e., the consistent condition) than non-associated stimuli (i.e., the inconsistent condition.). This, in essence, is the eponymous IAT effect. Typically, the reaction time differential, rather than the response accuracy differential across the two conditions, is used as a measure of differences in the strength of stimulus associations. Social-cognitivists use the existence of these IAT effects to draw inferences about unconscious cognitive activity or hidden prejudices and beliefs.

In the seminal IAT paper described above and in subsequent papers (e.g. Greenwald 2003) the authors suggested a model of implicit attitudes in which IAT effects reflect associations in memory that are tacitly theorised to be the building blocks of stereotypes or attitudes. However, their conception of associations is "theory uncommitted" (Greenwald, Nosek, Banaji & Klauer, 2005) with regard to the structure of mental associations. This approach has provoked concern from others in the field (e.g. Blanton & Jaccard, 2006, De Houwer, 2009; Gawronski et al., 2009), who question the usefulness of the IAT in the absence of a working theory with regard to *what* the IAT measures. The common theme running between the critiques outlined by these commentators is that the theoretical construct of "implicit attitudes" is poorly defined, and this theoretical uncertainly has led to an ongoing debate with multiple possible interpretations of IAT and implicit measure data with regard to the underlying causal variables. A complete summary of this debate would be beyond the scope of this thesis, and of limited use to the current work. Because the current research is based on a behaviour analytic foundation, it rejects the use of mentalisms as explanatory constructs – and it is these varying casual roles of these constructs and hypothetical mental processes that the participant of the "IAT debate". What is worth noting, however, is that no associationistic, representational model of implicit attitudes has proved satisfactory in its validity or coherence with observed data, even to researchers who base their work on an associationist or mentalistic foundation. The new conceptual and methodological tools provided by the RFT approach may lead to a more parsimonious, empirically supported, and better functionally understood account of attitudes, both implicit and explicit, and also to similarly improved procedures for detecting the verbal histories underlying attitudes.

A Functional Approach to Implicit Attitudes?

As indicated thus far, concern has been expressed with regard to the poor understanding of the IAT on a functional, process level. Even one of its leading proponents (Gawronski et al., 2007), has characterised this as a "deplorable disconnect between basic research on the mechanisms underlying implicit measures and the somewhat wider reception of research using these methods."

Jan De Houwer (2006) has questioned the widespread use of the term "implicit measure" by many researchers in the absence of any attempt to define its use properly. The word "measure" can be used to refer to either to the objective procedure used to gain a measurement of some variable, or the outcome of such a procedure, in the sense that the particular numerical variable is considered an index of the construct to be measured. The common usage of "implicit measure" described in the paper does not refer to the procedure used to attain such a result, but rather certain functional properties of the measurement outcome. In relation to the IAT, these properties are that the IAT purportedly measures attitudes such that participants; a) are not aware of what is being tested; b) do not have conscious access to the attitudes being tested; and/or c) have no control over the measurement outcome. Interestingly, in line with a behavioural critique of the IAT and associated mentalistic assumptions, De Houwer (2009) emphasizes that it makes no sense to refer to a measure as "implicit" if one is not explicit about the functional properties of the measure that make it so, and that one must have empirical research to back up that claim.

Leading from this assertion, De Houwer (2009) criticizes the IAT for its relative lack of empirical research into those functional properties. A single study (Monteith, Voils, & Ashburn-Nardo., 2001) has suggested that many participants are aware of what a given IAT is examining, which calls into question the "implicitness" of the IAT in that sense. In the study, participants showed strong "implicit" racial biases on an IAT. The researchers then examined the participants' responses to this bias, and discovered that the majority of participants had become aware of the bias the IAT was probing for. Many participants attributed their IAT bias to race related factors and expressed guilt at such a bias. Interestingly, participants who displayed small IAT biases were less likely to attribute their bias to race related factors, indicating that those who read as most strongly biased with an IAT are the most capable of explicitly relating the result to factors based on race.

In a functional, RFT driven approach, the IAT is viewed as a measure of an individual's verbal history and practices, which may or may not in turn reflect personal attitudes or affective states and dispositions. IAT effects are conceived in terms of participants' fluency with the relevant verbal categories and their degree of experience at juxtaposing members of those verbal categories. For instance, an individual who has many dealings with people of a specific race, and has encountered both pleasant and unpleasant individuals from this racial group, will likely find it easy to juxtapose racial and evaluative terms in an IAT according to the test rules across the two test blocks. Such an individual will show no IAT effect (i.e. response time or accuracy differential across the text blocks). On the other hand, if they have experienced mostly unpleasant individuals from one racial group or other, the juxtaposition of response rules across the IAT blocks will likely expose a fluency differential across those two blocks (i.e., an IAT effect). This is the behavioural model of the IAT (see Roche et al., 2005).

The behavioural model of the IAT was tested empirically by Gavin, Roche & Ruiz (2008) using nonsense syllables as stimuli and experimentally produced derived relations between them as laboratory analogs of verbal relations between words in the vernacular. Two equivalence relations were established in the usual manner, leading to the two classes of nonsense syllables, labelled here as A1-B1-C1 and A2-B2-C2, where the A-C relations were derived, not reinforced. An IAT-type test was then administered to measure participants' ability to respond in the same way to common class member pairs (e.g., A1 and C1) compared to cross-class pairs (e.g., A1 and C2). Not surprisingly, more errors were made in responding under rule conditions in which a common response was required for incompatible, compared to compatible stimuli. Thus, IAT effects were generated using only directly manipulated variables, without recourse to mentalistic language or appealing to hypothetical processes. These entirely laboratory produced IAT effects were subsequently shown to be manipulable vis-à-vis reversals of some of the baseline relations underlying the derived equivalence relations (Ridgeway, Roche, Gavin & Ruiz, 2010). Such findings strengthen any claims that the IAT test format is sensitive to a participant's history of relating the test stimuli, perhaps even including those that a participant would wish to conceal. However, they pose a challenge to any view that IAT effects are necessarily a reflection of internal beliefs, intentions, or predispositions. This may be the case, but these are not necessary conditions for IAT effects to emerge, and thus the scientifically conservative position to adopt on the matter is a functional one. What we can say with certainty is that IAT measures past stimulus relations in the history of the test taker, and that these stimulus relations may be directly established or derived.

The behavioural model of the IAT is critical of several features of the IAT presentation format and it is important that we briefly outline these concerns as they will be addressed in this thesis through the development of our new test methodology. Firstly, in the experimental analysis of behaviour, reaction times alone are not typically treated as indicative of the strength or stability of any instance of behaviour. Instead, response accuracy or fluency (combined accuracy and speed) measures are usually used for this purpose. A move to fluency as the primary measure over response time also circumvents several problems regarding the manner in which response latencies (i.e the time between stimulus presentation and the participant emitting a correct response) are calculated using the IAT (see Ridgeway et al., 2010 for a detailed critique).

Secondly, the putative feedback presented during the IAT may in fact function as a form of punishment. Participants are informed only when erroneous responses are produced, by the presentation of a red X on screen. No feedback is provided following correct responses. This is a less than efficient way to establish response fluency and the delivery of punishment has unknown affects on subsequent responses and response rates. This imbalanced feedback procedure makes the emergence of IAT effects more likely, but for reasons not clarified by its inventors. More specifically, this procedure ensures that whichever of the two test blocks contains less errors (the relationally inconsistent block) will also be the block that involves less punishment, therefore possibly exaggerating the accuracy and reaction time differences across blocks in the expected direction. This exaggeration may occur because rapid responding in the consistent block is negatively reinforced by the removal of both interruptions to response fluency and the omission or reduction of negative feedback. In contrast, rapid responding in the inconsistent block is punished, thereby leading to slower responding. Previous research has shown that response caution during difficult IAT tasks can partially explain IAT effects (Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007). The author would suggest that imbalanced and negative feedback during the IAT produces precisely such imbalanced response caution.

Thirdly, the most popular IAT scoring technique (D-algorithm; Greenwald, Nosek & Banaji, 2003) involves a standardized reaction times measure (taken across a fixed trial block) as the core index of association strength - which we would functionally define as the strength of the participants' reinforcement history vis a vis the relations of interest. However, the trajectories of reaction times across trials are not assessed and response rates and changes in rate are usually irrelevant to the measures (see also Blanton & Jaccard, 2006). In line with behaviour analytic tradition, learning *rates* to predetermined criteria may be the better primary behavioural measure over standardized latency scores.

Fourthly, the IAT's scoring method involves data cut off points and participant elimination procedures designed to stabilize data and increase statistical significance in subsequent analyses. Thus, the scoring method involves the creation of data stability through means other than improved stimulus control. In contrast, the stability of participants' behaviour during the FAST test will be enhanced across research studies using methods designed to improve experimental control over response variability, the analysis of which is an important part of behaviour analytic participant matter (Sidman, 1960).

Interestingly, in line with this functional perspective, some social cognitivists have suggested an environmental-style account of the IAT. Specifically, it was first suggested by Karpinski & Hilton (2001) that IAT effects may reflect only the word and concept associations to which a person has been exposed to in their past. They may not reflect the extent to which the person endorses those evaluative associations. Other researchers have suggested that participants' experiences of the words employed in an IAT test can alter the IAT effect itself (see McFarland & Crouch, 2002; Ottaway, Hayden & Oakes, 2001; but see also Dasgupta, Greenwald & Banaji, 2003; Gawronski, 2002).

In recent years, one explicitly behaviour-analytic alternative to the IAT has been developed. The Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes, Hayden, Barnes-Holmes, & Stewart, 2006) is in many ways procedurally similar to the IAT. However, each trial of the IRAP displays two stimuli (rather than just one) on screen (e.g., “Child” and “Sexual”) along with a contextual cue which specifies the relation between the two stimuli (e.g., “Same” or “Opposite.”). The participant is required to respond quickly to this resulting relational statement (“Child SAME Sexual” or “Child OPPOSITE Sexual”) with a key press that corresponds to one of two response options (e.g. for “TRUE” press z, for “FALSE” press m). Feedback is only presented if the response in incorrect as defined by the block rules (a red X is displayed) or if the response latency is lower than the stated criterion (the words “too slow” are displayed). Like the IAT, trials in the IRAP are organized into blocks. In one type of block, the responses that produce the “correct” feedback are those that are *consistent* with social norms (e.g. Child OPPOSITE Sexual – TRUE = Correct) while the other block requires responses that are *inconsistent* with social norms. A child-sex IRAP would function as follows. In the consistent blocks participants would be required to respond to the statement Child SAME Sexual as FALSE and Child OPPOSITE Sexual as TRUE. In the inconsistent blocks, the participant would be required to respond to the statement Child SAME Sexual as TRUE and to Child OPPOSITE Sexual as FALSE. The core assumption of the IRAP is that participants will respond more quickly to relations that are consistent with their verbal and nonverbal history with the stimuli.

The behavioural framework within which IRAP results are interpreted is the Relational Elaboration Coherence model (REC; Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010), informed by Relational Frame Theory. According to the REC model, each individual trial on the IRAP produces an immediate and brief response to the relation presented before the participant presses a response key. The probability of this initial response is a function of participants’ verbal and non-verbal history with the stimuli and the current contextual cues (i.e., the relational stimulus, such as the word *opposite* on screen). The most probable response will likely be emitted first, and as such, if this immediate response coheres with the response required by the current IRAP trial rule, then the response latency will be lower. If the required response is in opposition with the participants' immediate relational response, then the correct (as defined by the contingencies of the specific trial) response will be emitted more slowly. Across multiple trials, the average latency on inconsistent trials will be higher than the average latency for consistent trials.

The foregoing provides a basic explanation for the IRAP effect, but the REC model extends the explanation to account for why implicit measures and explicit questionnaire methods so often diverge in their results. More specifically, when completing questionnaires or other so called "explicit" measures of attitudes, the participant is under little time pressure and can therefore engage in complex and extended relational responding (i.e., thinking) which allows them to produce a response which is coherent with other responses in their behavioural repertoire (see Barnes-Holmes, Hayes & Dymond, 2001) such as; "it is wrong to categorize children as sexual." It is also possible under these circumstances to produce a response which coheres with the social expectations of others. However, when exposed to the IRAP procedure, participants are under significant pressure to respond quickly (commonly within 2000 ms) and, therefore have little time to engage in the elaborate relational responding processes necessary to produce alternative socially desirable responses. In effect, the most likely responses under time constraint conditions are those that are immediate and brief and therefore direct measures of history, unmediated by local relational activity.

One IRAP study set out to determine if the IRAP could be used to differentiate between child sexual offenders and a normal population group (Dawson et al., 2009). Sixteen participants who had been convicted of a contact sexual offence against a child (the offender group) and sixteen male non-offenders recruited from a college population (the control group) completed an IRAP procedure and a Cognitive Distortion Scale (CDS; Gannon, 2009). The IRAP stimuli consisted of two category labels ("Child" and "Adult") and two sets of target stimuli (“sexual” words and “non-sexual” words). During the consistent blocks, participants were required to respond with “true” to the relations “Adult – Sexual” and “Child – Nonsexual” while in the inconsistent blocks participants were required to respond in the opposite way.

The IRAP successfully detected a difference between the control group and the offender group. Furthermore, the IRAP was able to identify the specific relation on which the two groups differed. On Child-Sexual trials, the offender group did not show a significant IRAP effect, responding equally quickly to Child-Sexual-False and Child-Sexual-True. According to the REC model this effect would be due to the offender group immediately responding to child-sexual relations as both *true* and *false*. This makes sense in the context of the Implicit Theories of Sexual Offending Model (Gannon, 2006) which states that child sexual offenders conceive of children as simultaneously being sexually receptive, while also being innocent and non-threatening when compared to potential adult sexual partners. In the Dawson et al. (2009) study, the Cognitive Distortion Scale (CDS) did not differentiate between the two groups. Again this may not be surprising. In this case, all members of the offender group had participated in a Sexual Offender Treatment Program, which targeted the cognitive distortion associated with offending. When under low time pressure during completion of the CDS, the offenders would have been able to respond consciously in a manner which was consistent with their treatment, thus not showing any difference from controls.

The preceding work has provided a powerful functional account of “implicit attitudes”, and has garnered attention from the same researchers who were critiquing the IAT on a similar basis. De Houwer (2011) has outlined the problems inherent in using behavioural manifestations (e.g., response latencies on IAT tasks) of assumed cognitive structures (e.g., unconscious bias) to infer the existence of those very constructs. In the same paper, he outlines the potential benefits for both cognitively oriented and functionally oriented (e.g., behaviour-analytic) research in combining both approaches into a functional cognitive framework. It is argued that the functional approach is useful to the social-cognitivist, for example, in that it provides useful information about the environmental causes of behaviour and the environmental variables that can be experimentally manipulated to produce or alter a behavioural effect without necessary reference to mental constructs as causal events. This information, in turn, allows the cognitively oriented psychologist to make more informed inferences about the mental constructs assumed to mediate such behavioural effects by eliminating *a priori* assumptions and providing clear information about the input to mental processes. The functional and cognitive approaches can be integrated but remain conceptually distinct. However, the functional approach informs the cognitive approach as to the facts which need to be accounted for with mental explanations, without reference to the mental explanations themselves, while the cognitive explanations provide a stable theoretical framework for the development of functional knowledge.

The work in this thesis adopts a functional approach to implicit testing but remains appreciative of the need to remain mindful of the heuristic value of mentalistic organising concepts. Mental explanations can sometimes benefit the functional analyst by organising existing knowledge and making useful predictions which can prove fertile to future research even if that research then takes place using entirely non-mentalistic concepts. For example, the concept of attitude is intrinsically mentalistic, and yet can serve as a useful metaphor in organising specific varieties of verbal behaviour (evaluative verbal responses) for specific research purposes (e.g., predicting future nonverbal behaviour such as racially motivated violence). However, the construct of "implicit attitudes" as used by cognitivist researchers is poorly defined and of limited use except to orient the research towards a possible phenomenon of interest.

A behavioural definition of attitudes has already been provided (see above), which leaves the concept of "implicitness" in need of a functional definition. De Houwer and Moors (2007) has provided an extensive analysis of what could be meant by the term "implicit" that could serve to guide functional research. The authors identify the term "implicit" with the term "automatic" – that is, a process is implicit if it is "*a measurement outcome that reflects the to-be-measured construct by virtue of processes that have the features of automatic processes".* Automatic processes are characterised by being unintentional, uncontrollable, unconscious, effortless, and fast. However, automaticity is not an all or nothing definition – the manner in which a process is automatic is dependant on "enabling features" in the measurement procedure, and the authors suggest that a "decompostional, feature based" approach may shed the most light on automatic/implicit processes. This suggests that a functional approach is of the most benefit in this instance. From the functional perspective of this thesis, the notion of implicitness can be conceived of as referring to the fact that the contingencies controlling the relevant responding are not discriminable by the participant, or that the participant will not (or cannot) overtly tact the verbal contingencies operating on their behaviour.

With the continued interest in "implicit testing" style methodologies across all fields of research in psychology, the need for functionally understood measures only grows. With the emerging analyses of the popular IAT methodology in behavioural terms (Gavin et al., 2008) and the key role played by stimulus equivalence in language and attitudes, the time is ripe for a return to the core processes leveraged by the Watt et al. (1991) – that a participants' social history interferes with the acquisition of new relating behaviours. Building from these behavioural first principles, the current research developed a subtle, indirect testing methodology (which would be described as *implicit* in the social cognitive literature) that can be functionally understood, easily utilized, and quickly administered by researchers in any field of psychological study.

In seeking to develop such a measure, we return now to the Watt et al. (1991) methodology described above. This methodology was built entirely upon well understood behavioural processes – empirical support for which has only grown in the past twenty years. However, while the Watt et al. procedure has proven to be of considerable interest to behaviour analysts it has not been adopted widely as a testing methodology in the world outside the laboratory or by psychologists in mainstream psychology, as have several popular “implicit” tests, such as the Implicit Association Test (Greenwald et al., 1998). I offer three main reasons why this is the case. Firstly, the Watt et al. stimulus equivalence paradigm requires an understanding of the stimulus equivalence phenomenon and the conditional discrimination training methodology with which it is usually associated. Thus, it is not easily accessible to non-behaviour analysts. Secondly, the test format is very demanding of participants insofar as it requires considerable attention and motivation in order to complete training and testing phases of the procedure. Finally, the test format requires considerable time compared to other popular “implicit” tests, such as the IAT. While the latter typically require only a few minutes, completing the stimulus equivalence methodology requires approximately 20 minutes.

Another potential shortcoming of the Watt et al. procedure is one that may compromise its subtlety. Specifically, in this procedure, two incompatible stimuli are presented together as sample and comparison, respectively. While a participant has no way of being certain what the purpose of the measure is, the fact remains that the stimuli whose relation is under analysis are presented simultaneously during the testing phase. Thus, *procedural implicitness* is compromised by the Watt et al. technique. Procedural implicitness refers to the degree to which the relations under analysis and the purpose of the test are discriminable by a participant. This is to be distinguished from *outcome implicitness*, which refers to the implicitness of the stimulus associations being measured (i.e., whether the participant has ever discriminated the associations in the past; see De Houwer, 2006).

An important but simple modification can retain the basic core process of the Watt et al. methodology while at the same time disguising the purpose of the test more fully. More specifically, if we attempt to establish a functional response class, instead of a derived equivalence relation, an entirely novel test format presents itself. That is, instead of presenting the two stimuli of interest simultaneously to assess the probability of a matching response during an equivalence test, it is possible to present the stimuli individually on separate trials and attempt to establish a functional stimulus class containing them. This can be done by establishing distinct response functions for stimuli that are suspected of being related as a result of prior social interaction. The rate at which a participant learns to produce the common response for both of these stimuli, presented separately, can be compared to the learning rate for producing distinct responses for these two stimuli, again across separate trials. Such a procedure requires no conditional discrimination training, is not demanding on the participant (i.e., there are no relations to derive), and is fast to administer.

As an example of the foregoing, we might assess whether or not the words *Mathematics* and *fun* are related for a sample of school children with difficulty learning maths. We could present them with each of four stimuli, individually and in a random order and across several re-cycled blocks of trials. The stimuli should consist of the words of interest (*mathematics* and *fun*) and also two novel and entirely unrelated words; nonsense syllables may work best. Upon the presentation of each stimulus, the participant may be required to produce one of two responses. These responses may be positional on a computer keyboard, or may involve clicking on a specific discriminated response key on a computer screen. Correct responses can be easily reinforced verbally or with tokens. Importantly, the words of interest share a common response, while the novel control stimuli (e.g., nonsense syllables) also share the remaining response function. The number of trials required for the participant to reach a preset fluency criterion in this block of trials, represents an index of the pre-existing strength of the relation between the words *mathematics* and *fun*.

We may also re-run the test block, but with the important difference that now we establish a response class for the word *mathematics* and one of the control words. Similarly, the word *fun* and the remaining control word now share a response. If the words *mathematics* and *fun* have been related in the history of the participant, then this latter task block will require a larger number of training trials than the former block described above, simply because the latter task is inconsistent with the social history of the participant (i.e., the functional class is inconsistent with the socially established functional or equivalence class containing the words *mathematics* and *fun*). A baseline rate of acquisition of response classes using novel and arbitrarily chosen stimuli may also be recorded. We may then assess the extent of the facilitating or retarding effect of pre-existing stimulus relations between *mathematics* and *fun* on learning during the two test blocks. Taken together, this approach to assessing the strength of a stimulus-stimulus relation can be referred to as a function acquisition speed test (FAST).

The first empirical chapter of this thesis aims to develop the new methodology and test its utility in detecting stimulus relations which have been established in the laboratory. This will provide an important proof of concept for the methodology, demonstrating its ability to detect relations over which we have full experimental control. The experiments in thus chapter created a known history of stimulus relations which allows the experimenter to determine if the process was successful in detecting said history without concern about issues of stimulus selection such as familiarity and salience which often obfuscate results when procedures are tested with “real world” stimuli. This investigation is conducted across four experiments. Experiment 1 trained participants in simple A-B relations between nonsense stimuli before exposing them to the novel test format (dubbed the Function Acquisition Speed Test; FAST) designed to detect those same relations. The test format proved able to detect the experimentally established history in the majority of participants. Experiment 2 trained simple functional, affective relations between nonsense syllables and picture stimuli before exposure to the FAST. This experiment did not prove successful, potentially due to confounds caused by different stimulus types (words and pictures) utilized in the FAST. These issues will be returned to in later experiments, while the remainder of the experiments described in Chapter 1 continued to use only nonsense syllable stimuli.

Experiments 3, 4 and 5 built upon the success of Experiment 1 by creating more complex or more robust histories which incorporated derived relational responding under different conditions. Establishing the FAST's utility in detected histories of derived relational responding is essential due to the instrumental nature of this type of responding in complex human behaviour and attitudes in particular. Experiment 3 trained linear A-B-C relations before testing for derived A-C relations. Participants were then exposed to a FAST procedure to probe for the derived A-C relations. Again, the procedure was successful in detecting these relations in a majority of participants. In Experiment 4, in order to make certain that those participants who did not show the predicted FAST performance did so due to insufficient training (rather than a deficiency in the procedure itself) Experiment 4 exposed participants to an extensive, intensive regimen of equivalence training over a period of 10 stages and 3 days before they completed a FAST. This Experiment was a repeated single participant design. The FAST successfully detected the robust equivalence training history in 5 out of 6 participants, providing a firm foundation to extend the FAST research into a real world context. Experiment 5 once again trained A-B-C relations. In this experiment, half of the participants were exposed to a FAST *before* the derived relations were tested for, while the other half were exposed to the FAST *after* an equivalence test. For participants who had not been exposed to equivalence testing, the FAST was unable to detect the history of stimulus equivalence training, indicating that untacted relations may not be readily detected by the FAST procedure.

In the second empirical chapter, a number of experiments were conducted to examine then effect of small procedural modifications. In each of these experiments, participants were trained in simple A-B relations, and then exposed to a series of FAST procedures. In each experiment, there were multiple experimental groups. Experiment 6 explored the effect of different response windows on FAST performance between groups. A response window of 4s produced the worst outcomes, while the windows of 1s, 2s, and 3s produced similarly effective but functionally different patterns of results. Experiment 7 varied the number of practice blocks between participants, but no statistically significant differences were found. Experiment 8 investigates the effect of varying the fluency criterion for completing a FAST block. The results suggested that a balance between very rigorous and very lax criterion produced the most reliable results.

Building upon the strong foundation of the first chapter, the third empirical chapter involved participants completing FAST procedures designed to probe for real world verbal histories. In Experiment 9, the FAST procedure used real words as stimuli, with the words *spiders* and *disgust* as the stimuli of interest in the FAST test blocks. Participants also completed pen and paper measures measuring Fear of Spiders and Social Desirability. On a group level, the sample showed a strong history of relating the target stimuli. The explicit measures showed an overall low rate of spider fear, and neither the FAST nor the explicit spider fear questionnaire interacted with Social Desirability. Experiment 10 used a similar procedure, but instead used the words "*immigrant"* and "*cheat"* as the target stimuli in the FAST test blocks, and used explicit measures which detected prejudiced attitudes towards immigrants. The sample showed no significant FAST effect, indicating a lack of relating the two stimuli. However, participants with high social desirability scores tended to demonstrate larger FAST effects. The final experiment in Chapter 3, Experiment 11, used groups of images as stimuli rather than words. Participants completed three FAST tests, in which the target stimuli were Tanner images of young females and erotic images. The age of the female images varied from FAST to FAST. Participants demonstrated a strong negative FAST effect (i.e. suggestive of a history in which the stimuli were in a frame of opposition) when the female stimuli were at Tanner Age 3 (11-13), but this effect quickly began to reverse as the apparent age of the female images increased.

Chapter 2

Testing the FAST under laboratory controlled conditions

The aim of this thesis is the development of a behaviour analytic “implicit” test for analyzing histories of relational responding and stimulus equivalence. The previous chapter documented the background of implicit testing in general, and the history of behaviour analytic work to develop (Watt et al., 1991) and functionally understand (Gavin et al., 2008) similar tests. Synthesizing the approaches, insights, and ideas from existing research in the field suggests a new test format, which will be detailed and tested in the current study.

The Watt et al. (1991) procedure demonstrated the potential for a behaviour analytic test methodology based on the finding that a participant’s pre-experimental social and verbal history interferes with the formation of new stimulus equivalence relations. Laboratory studies (Hall, Mitchell, Graham, & Lavis, 2003; Roche et al., 1997, Tyndall, Roche, & James, 2004, 2009) corroborated the original Watt et al. research in laboratory controlled conditions and confirmed that incongruous stimulus relations impede the training of new functional or stimulus equivalence relations.

The foregoing suggests a test format in which a participant is required to form a stimulus class across two different test blocks. However, requiring participants to form novel stimulus classes makes it necessary to present the stimuli of interest simultaneously. The key innovation which informs the FAST procedure is the fact that the format requires participants to learn response functions for four stimuli, two of which are the target stimuli of interest, and two of which are neutral and unrelated stimuli.

In one block, the reinforcement contingencies would require the same functional response to each of the test stimuli, while reinforcing a second response to the other two stimuli. If the test stimuli are related to one another in the participant's history, then this task will be relative easy. In the other block, the reinforcement contingencies would support the emergence of response functions that are incongruous with those established in the participants' pre-experimental history – different responses would be required to each of the test stimuli (i.e. the contingencies would act as a behavioural disrupter). In contrast to similar testing formats (e.g. the IAT and the IRAP), response latency would not be the datum of interest. Rather, the rate at which the participant learns to respond fluently to each set of contingencies (i.e., learning rate) would be the dependent measure of interest. The longer it takes a participant to learn to respond fluently under the control of the incongruent reinforcement contingencies, the more well-established the target relations must have been in the participants' history.

The experiments described in this chapter apply the FAST procedure to detect relations that are under the control of the experimenter. By way of a number of pre-experimental training phases (which vary across experiments), the experimenter establishes a known history of relating two or more nonsense syllable stimuli. The developing FAST procedure is them tested for its ability to probe for the experimentally established histories.

In Experiment 1, a simple A-B relation was trained between two nonsense syllables. After this training procedure, participants were exposed to a FAST procedure designed to detect the newly trained A-B relations. In Experiment 2, two functional relations were trained (A1 – erotic, B1 – aversive), and a FAST procedure was then employed to detect the A1 – erotic relation.

Experiment 3 exposed participants to a linear matching to sample procedure in order to form two three member equivalence classes from nonsense syllables (A1-B1-C1 and A2-B2-C2). The A and C stimuli were not directly paired, resulting in a derived A1-C1 relation which was then probed for with a FAST procedure. Experiment 4 similarly focused on using the FAST to detect derived relations between nonsense syllables. However, Experiment 4 utilized a more rigorous and extensive training procedure to train the two three member equivalence classes. Participants completed nine successive stages of Matching to Sample training. The MTS training used a One-To-Many paradigm, and required an observation response to the sample stimulus before the comparison stimuli appeared. As the stages progressed, the criteria for completion of each block increased, as did the delay between observation response and the display of comparison stimuli. Further, in the final three stages, reinforcement (via corrective feedback) was reduced.

In Experiment 5, participants were again trained using Matching to Sample to form two three member equivalence classes. In this experiment, half of the participants completed a FAST *before* the derived relations were tested for, in order to test the FAST's ability to detect derived relations that have not yet been tacted by the participant.

Experiment 1: Using the FAST to detect A-B stimulus relations established in the laboratory.

In the current experiment, the nascent FAST procedure will be tested in the most elementary manner possible. Before tackling real word stimulus relations, or even derived relations, the FAST must prove capable of detecting the simplest possible stimulus-stimulus relation.

A brief pre-experimental training procedure will establish a known history of relating two nonsense syllables. It is predicted that this history will prove sufficient that participants will have greater difficulty in learning a pair functional responses such that the two related stimuli do not share a function (inconsistent block) as compared to a block in which participants are reinforced for the same functional response to both related stimuli (consistent block). In other words, this experiment serves as the most elementary proof of concept for the FAST procedure.

Method

Participants

Twenty-one volunteers participated in the study as participants. Three were eliminated due to failing to reach performance criteria during the FAST test or the baseline blocks (see Results). Of the eighteen remaining volunteers, six were male and 12 were female. Ages ranged from 19-36 years (M = 22.94, SD = 3.72). The participants were recruited both from acquaintances of the experimenter and using a 'snowballing' technique whereby participants were asked to recruit a further participant for participation. Informed consent was obtained in writing from all participants.

Apparatus and Stimuli

All phases of the experiment were presented to participants on an *Apple Macbook*® laptop computer with a 13-inch monitor (1024 X 768 pixel resolution). Stimulus presentations were controlled by the software package *Psyscope* (Cohen, MacWhinney, Flatt & Provost, 1993) that also recorded all responses. Stimuli consisted of fourteen nonsense syllables (ter, lar, jum, mip, ler, mau, hox, yun, pim, kon, geq, kav, zuv, jin) randomly assigned to their roles as samples, comparisons and FAST test stimuli. These will be referred to using alphanumerics in future.

Ethics

All participants were presented with and signed a consent form before proceeding to first phase of the experiment (See Appendix 1). Participants were told informally that performance on the task would not allow the researcher to make any individual psychological assessments but may allow for group patterns to be identified. All participants were informed of the true nature of the study after participation and were given the opportunity to withdraw their data at that stage. This study, and those that followed, were approved by the National University of Ireland, Maynooth Ethics Board prior to beginning data collection.

General Experimental Procedure

The experiment consisted of two phases. Phase 1 (stimulus matching procedure) required approximately 8 minutes to complete. Phase 2 (FAST) consisted of four FAST blocks – two baseline blocks, a consistent test block, and an inconsistent test block. Each FAST block typically took 3-6 minutes to complete. All phases were presented consecutively by the computer software, which also controlled the delivery of instructions at the beginning of each phase. Participants sat comfortably at a standard computer desk and viewed the computer screen at a distance of approximately 60-70cm and at eye level.

Phase 1: Matching-To-Sample Training. In this first Phase, participants were exposed to four matching-to-sample training tasks, each designed to establish two simple stimulus relations. The relations trained were; choose B1, not B2 when A1 is present; choose B2, not B1 when A2 is present, choose A1 not A2, when B1 is present, and choose A2 not A1 when B2 is present (see Figure 1). The purpose of this phase was to establish two laboratory controlled stimulus relations that could be employed to assess the utility of the FAST in determining the existence and strength of known stimulus relations.

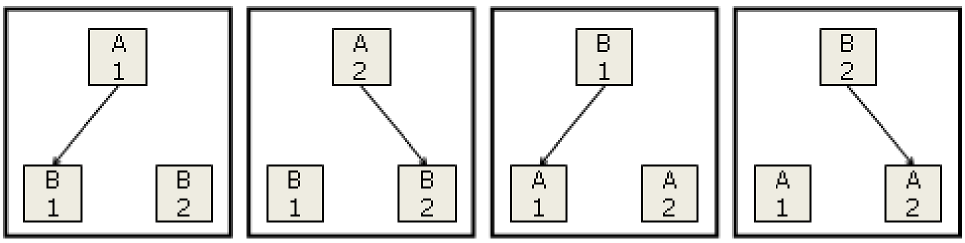


Figure 1. The stimulus relations trained during Phase 1. Solid lines indicate reinforced matching responses.

Participants were presented with the following instructions at the onset of Phase 1.

In a moment some words will appear on this screen. Your task is to look at the word at the top of the screen and choose one of the two words at the bottom of the screen by "clicking on it" using the computer mouse and cursor. During this stage the computer will provide you with feedback on your performance. You should try to get as many answers correct as possible. If you have any questions please ask them now. When you are ready please click the mouse button.

All trials were presented on the computer screen against a white background. A trial began with the presentation of the sample stimulus at the top centre of the screen in black 24-point font. One second later, the two comparison stimuli were displayed in black 24-point font in the bottom left and right corners of the screen (see Figure 1). The positions of the comparison stimuli were counterbalanced across trials. All stimuli remained onscreen until the participant responded by clicking on one of the comparison stimuli using the computer mouse pad and cursor. Immediately upon a response, the screen cleared and corrective feedback ("Correct" or "Wrong" in red 48 point font) appeared in the centre of the screen for 1.5s.

Trials were presented in blocks of 32 (i.e., each of the four trials presented 8 times). Participants were required to complete successive training blocks (up to a maximum of four cycles) until a criterion of 31 or 32 correct responses in a single 32-trial block was met (i.e. minimum 96.9% correct).

Phase 2: Function Acquisition Speed Test (FAST)

General FAST Block procedure. Each FAST block utilises four stimuli. In each block, participants are required to learn a common response to one pair of stimuli and a different response to the other pair (e.g. press "a" for X1 and Y1, press "j" for X2 and Y2). The block continues until a participant reaches a predetermined criterion (10 correct responses in a row). The number of trials a participant requires to reach this criterion on each block is the primary datum of the FAST.

All trials were presented on the computer screen with a white background. A trial began with the presentation of one of the four nonsense syllable stimuli (e.g. X1, X2, Y1, or Y2) in the centre of the screen in 48-point black font. The stimuli remained on-screen for a period of 3s or until a response was emitted (i.e., a 3s response window was enforced). Each of the four stimuli was presented in a quasi-random order in blocks of four trials (i.e. consecutive exposures to any one stimulus were not possible).

Immediately upon the production of a response, corrective feedback was presented (i.e. either “Correct” or “Wrong” in red 48-point font in the centre of the screen for 1.5s). If no response was emitted within the 3s response window, an incorrect response was recorded but no feedback was provided. In that case, the screen cleared and the next trial began immediately upon the end of the 3s response window. Participants were exposed to trials until a criterion of 10 correct in a row was reached. That is to say, participants were required to produce ten correct responses across any contiguous sequence of 10 trials. If the participant reached this criterion, then the block ended automatically and the instructions page for the next block would be presented.

A predetermined limit of 100 trials was enforced because pilot research had indicated that once this limit was reached, the participant was unlikely to complete the block before giving up or being asked to cease by the experimenter.

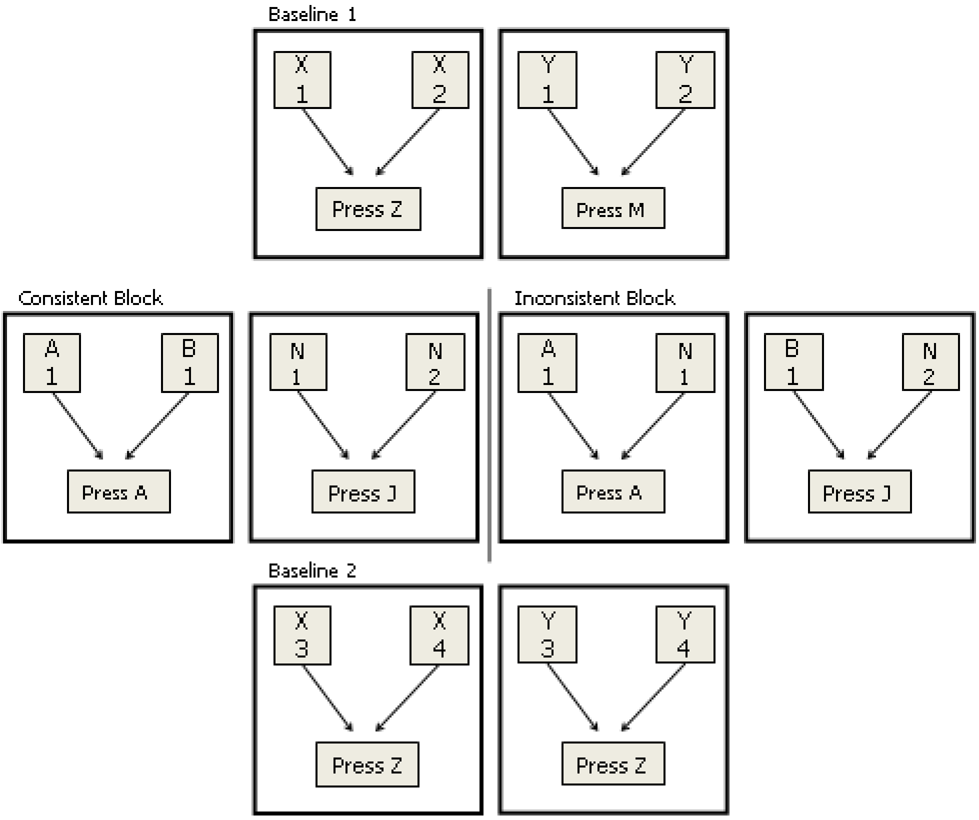


Figure 2. A schematic of the various response functions assigned to stimuli during baseline and FAST blocks. The three panels, read from top to bottom, also represent the sequence of presentation of the two baselines blocks and the two FAST blocks.

Baseline Blocks. The purpose of the baseline blocks was to establish a baseline level of response class acquisition using novel and previously unrelated stimuli, against which acquisition rates with target stimuli could be compared. The baseline blocks each involved novel and unique nonsense syllable stimuli with which the participants had no previous experience. Baseline 1 employed X1, X2, Y1, Y2 as stimuli, while Baseline 2 employed X3, X4, Y3, Y4 as stimuli. Baseline block 1 required participants to learn to “press left” when presented with X1 or Y1, and to “press right” for X2 and Y2. Similarly, baseline block 2 required common functions for X3/Y3 and X4/Y4.

Two baseline phases were presented (one before and one following the test blocks) in order to assess the stability of baseline rates of function acquisition across time. Administering two baseline phases also had the advantage that it would allow for the calculation of a mean baseline acquisition rate if baseline performances proved to be unstable across time (see Results). The following instructions were delivered at the start of each baseline phase:

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.

Test Blocks. In this phase, participants were exposed to the two-block FAST test using the A1 and B1 stimuli from Phase 1 and two further novel nonsense syllable stimuli, N1 and N2. One of these blocks (the consistent block) established two functional stimulus classes (A1-B1 and N1-N2) that were consistent with the stimulus-stimulus relations established in Phase 1 (i.e., in which A1 was matched with B1). The other block (inconsistent) established two functional stimulus classes (A1-N1 and B1-N2) that were inconsistent with the relations established in Phase 1. The order of the consistent and inconsistent blocks was randomized across participants.

Participants were presented with the following instructions at the onset of each FAST block.

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 A key or the J 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.

In summary, the FAST blocks attempted to establish two response classes under two conditions; one in which previously related stimuli participated in the same functional stimulus class, and one under which they participated in distinct functional stimulus classes. Participants were exposed to a series of blocks in the following sequence; Baseline 1, FAST (Inconsistent and Consistent blocks, in random order), Baseline 2 (See Figure 2). After the completion of each block, the instructions page for the next block appeared, allowing the participant to begin the next block whenever they were ready without interruption from the experimenter. Upon completion of the final baseline block, a page appeared thanking the participant for their participation and instructing them to contact the experimenter.

Results

Of the twenty-one volunteers who completed the experimental procedure, the performances of three participants were not included in the data analysis. Specifically, Participants 2 and 6 were eliminated on the basis that they failed to reach response fluency on one of the baseline blocks within 100 trials. Participant 8 failed to reach the fluency criterion on one of the FAST blocks (the consistent block).

Matching to Sample Training

All participants successfully completed Matching to Sample training. The average number of training trials needed to reach the criterion (97% correct) was 65.61. Table 1 illustrates the training requirements for each participant.

Table 1. The number of trials required to complete matching-to-sample training

Sub	**1**	**3**	**4**	**5**	**7**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
Trials	64	32	32	32	128	32	96	64	64	64	64	64	64	64	128	32	64	96

Baseline Blocks

Table 2 shows the number of trials required by each participant to reach the fluency criterion (10 correct responses in a row without error) for each baseline block. These data illustrate what appears to be large variation in performance from the first to the second baseline block. Upon close inspection, however, this variation would not appear to be consistent, and cannot therefore be considered a practice effect. A Wilcoxon signed ranks tests found the difference in trial requirements across the two baseline blocks to be non significant ($z = -1.198$, $p = 0.231$, two-tailed). Given the within-participant variability in performance across the baseline blocks, the mean number of trials to criterion across both baseline blocks for each participant was calculated. This figure was then used as the baseline acquisition rate against which to compare FAST block acquisition rates.

Table 2. The number of baseline trials required to reach the fluency criterion across both baselines

Participant	1	3	4	5	7	9	10	11	12	13	14	15	16	17	18	19	20	21
Baseline 1	26	84	12	41	40	27	42	70	24	88	20	48	66	21	35	49	20	12
Baseline 2	31	26	10	15	48	12	12	12	100	31	58	30	16	12	24	58	59	21
Mean	28.5	55	11	28	55	19.5	27	41	62	60	39	39	41	17	30	54	40	17

Function Acquisition Speed Test

Each participant completed two FAST test blocks; a consistent and an inconsistent block. Table 3 shows the number of trials to criterion for each test block and the order in which the FAST blocks were administered. It was expected that a larger number of trials would be required for participants to reach criterion on the inconsistent block compared to the consistent block. A Mann and Whitney U-test indicated that there was no significant order effect observed in the data. That is, trial requirement differentials across the FAST blocks were not significantly different *based on order* ($z = -0.442$, $p=0.67$, two-tailed).

Of the eighteen participants, thirteen showed a faster rate of response function acquisition in the consistent block, compared to the inconsistent block, as expected. Four participants showed a small acquisition rate differential in the unexpected direction, while one participant (S7) showed no acquisition differential across the FAST blocks. The mean differential in FAST block trial requirements in the expected direction was 27.1. In contrast, the mean differential in FAST block trial requirements in the unexpected direction was considerably smaller at -7.5, indicative of the expected FAST effect at the group level.

Table 3. The number of trials required to reach the fluency criterion across each of two blocks of the FAST test. The order in which the two test blocks were administered is also indicated (1 indicates that the inconsistent block was administered first, whereas 2 indicates that the consistent block was administered first).

Participant	**1**	**3**	**4**	**5**	**7**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
Inconsistent Block	88	60	40	20	22	25	32	15	43	40	10	76	23	47	32	100	22	22
Consistent Block	15	41	10	16	22	10	24	12	48	28	12	26	30	10	20	17	16	38
Difference	73	19	30	4	0	15	8	3	-5	12	-2	50	-7	37	12	83	6	-16
Order	1	2	1	2	1	2	2	2	2	2	1	1	2	1	1	2	1	1

A Wilcoxon Signed Ranks tests was performed on the acquisition rate data during the two FAST blocks. The acquisition rate differential across the FAST blocks was found to be significant ($z = -2.580$, $p = 0.005$, one-tailed), indicating a strong FAST effect at the group level.

In order to quantify the magnitude of each participant’s own fluency differential in response class acquisition across the FAST blocks, a simple *Strength of Relation (SoR) Index* was employed. This was calculated for each participant by subtracting the number of trials required to complete the consistent block of the FAST from the total number of trials required to complete the inconsistent block of the FAST, and dividing this differential by the mean number of trials required by that participant to reach criterion during the baseline blocks. In effect, the *SoR* index can be used here as a measure of the strength of the A1-B1 stimulus matching response, insofar as each participant’s acquisition rate differentials across FAST blocks is adjusted by a factor representing their own unique baseline acquisition rates for such task-types. Larger positive *SoR* indices indicate higher strength stimulus relations (i.e., A1-B1 stimulus matching responses are fast to acquire and are more resistant to disruption by competing reinforcement contingencies), while zero or negative responses indicate absent or reversed FAST effects (i.e., A1-B1 relations are slow to acquire and are more easily disrupted by competing reinforcement contingencies).

Table 4. Strength of Relation Indices [SoR]. Larger numbers indicate stronger A1-B1 stimulus matching responses (i.e., stimulus relations). Negative numbers indicate a reversal of the expected acquisition rate differential across FAST blocks (i.e., a reversed FAST effect).

Participant	**1**	**3**	**4**	**5**	**7**	**9**	**10**	**11**	**12**
[SoR]	2.56	0.35	2.73	0.14	0	0.77	0.3	0.07	-0.08
Participant	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
[SoR]	0.2	-0.01	1.28	-0.17	1.82	0.41	1.55	0.152	-0.01

Table 4 shows that for the majority of participants, the FAST test produced positive identifications of a history of A1-B1 training (i.e., it was sensitive to the participants' histories of relating A1 and B1). That is, positive *SoR* indices were calculated for 13 of the 18 participants. This table also shows that where *SoR* indices were negative, indicating a reversed acquisition rate differential across the FAST blocks, the magnitude of the index was small. In other words, the tendency is predominantly towards large positive *SoR* indices and small negative ones.

A one-sample Wilcoxon was also conducted to assess the significance of *SoR* index scores against zero. This analysis showed that *SoR* index scores differed significantly from zero in the expected positive direction ($z= -2.911$, $p = 0.002$, one-tailed), indicating a history of A1-B1 relational training for the participant group as a whole and a significant FAST effect.

The researchers then examined whether or not poorly established A-B relations during Phase 1 might partially account for those instances in which the predicted FAST effect was not observed. Some participants who failed to show a FAST effect had required a large number of training trials (128, Participant 7, SoR = 0), while others who failed to show a FAST effect had required a smaller number (64, Participants 12, 14, and 16). However, none of those who failed to show a FAST effect had completed the training in the minimal number of trials (32). It was suspected that requiring a larger number of training trials to complete simple conditional discrimination training may be indicative of poor stimulus control. In simple terms, the more training participants required to complete the MTS training, the more likely it was that fortuitous forms of stimulus control (or mere chance-level choices during matching-to sample tasks) controlled final performances. To check this idea, inferential correlational analyses were conducted to examine the relationship between the number of training trials required in Phase 1 and the final FAST effect. A Spearman's rho analysis indicated that training trial number was moderately negatively correlated with the raw trial requirement differential across the FAST blocks (r = - 0.495, n=18, p = 0.018, one-tailed). Training trial number was also found to be moderately negatively correlated with the *SoR* index of the FAST effect (r = - 0.439, n=18, p = 0.034, one-tailed). Thus, more rapid acquisition of the trained relations during Phase 1 was moderately associated with a larger FAST effect.

In summary, 13 out of 18 of the participants showed the expected FAST effect. At a group level, participants required significantly fewer training trials for consistent response classes to be established compared to inconsistent response classes. The *SoR* index also provided a statistically significant measure of pre-existing stimulus-stimulus relations at the group level. Finally, rapid acquisition of baseline relations during Phase 1 was found to be moderately associated with larger FAST effects.

Discussion

The current experiment established the basic proof-of-concept for the FAST methodology - that the procedure can successfully detect a history of specific stimulus-stimulus relations at the group level. Thirteen out of eighteen participants required more trials to complete the inconsistent block when compared to the consistent block, and learning during consistent block performances was faster than inconsistent block learning at the group level. Those participants who showed a negative SoR index (indicating a faster acquisition rate on inconsistent blocks) demonstrated only a small negative difference. The FAST either showed relatively strong effects in the predicted direction, or negligibly small differences in the opposite direction.

This experiment provides a strong foundation for the extension of the current experimental methodology into the investigation of other types of stimulus relations (such as functional relations and derived relations) and eventually into the detection of previously existing stimulus relations in the vernacular (i.e. real world verbal relations). However, there are a few issues that emerged from the current data which will require analysis in future experiments.

There was a noticeable level of variability in within participant baseline performances. Statistical analysis showed that the apparent decrease from Baseline 1 to Baseline 2 performances was not statistically significant. However, it did flag a potential difficulty going forward. The variation between baselines justifies the use of multiple baseline phases and of using the mean baseline to calculate our SoR index by demonstrating that a single baseline block score is not necessarily representative of a participant's baseline rate of learning during FAST blocks.

The foregoing also suggests the possibility that the same may be true of the FAST test blocks themselves. As such, future experiments will expose participants to multiple FAST blocks to assess the stability of FAST scores across time and exposures.

On a single participant level, the current results are promising. For 13 out of 18 participants, the FAST detected the history of relating A1 and B1. Rapid acquisition of the trained relations was moderately associated with positive SoR indices, which further suggests that the negative scores (which were mostly small and not significantly different from zero) were as a result of the trained relations not being effectively learned by the participant rather than as a failure of the procedure to detect the relations.

Another possible source of the negative SoR indices comes from the criterion set for the completion of a FAST block. Participants were required to respond fluently to the FAST task by responding correctly 10 times in a row while under time pressure (i.e. 3000ms response window). In cases where the target relations are weakly established, a single error on a FAST test block can have quite significant repercussions on the final datum. For example, a participant could have successfully learned the task, and be approaching the 10 in a row correct response criterion before a small motor control error caused by fatigue or a lapse in attention resets the response counter to zero. In effect, one error can require the entire sequence of 10 additional successively correct responses to be made, thereby adding at least 10 trials to the trial requirement for that block. The following experiment will examine the potentially disruptive effect of this strict criterion on test results by relaxing the criterion to allow for a single error within the sequence of 10 correct responses. The following experiment will also examine the utility of the FAST procedure in examining other types of stimulus relations that may be of interest to researchers using the FAST procedure in the world outside the laboratory.

Experiment 2: Using the FAST to detect functional relations established in the laboratory

Experiment 1 established that the FAST procedure is capable of detecting a history of simple stimulus-stimulus relations in participants. In order for the FAST to reach its goal of being a sensitive and reliable test for real-world histories of stimulus relations which might underlie attitudes of interest, it is necessary to investigate the procedure's ability to detect different types of stimulus relations that may typify stimulus relations germane to what we might call "attitudes". In the cognitive psychological attitude literature, attitudes are often conceived of as being composed of both a cognitive component and an affective component. The cognitive component of an attitude is seen to be based on beliefs, judgements and thoughts about the attitude object, while the affective component is based on feelings, emotions, and drives (McGuire, 1969). The extent to which an attitude or preference derives from once source or the other varies from attitude to attitude (Zajonc & Markus, 1982) and is often determined by the context in which the attitude came about (Edwards, 1990).

From within a behaviour-analytic framework, this distinction between "cognitive" and "affective" based attitudes can be conceived of as differentiating between attitudes which come about largely due to socially mediated verbal behaviour (i.e. cognitive based attitudes) or due to more automatic behaviour processes such as classical conditioning and/or evaluative conditioning. Both phobias and sexual fetishes serve as examples of what cognitive psychologists would describe as "affect-based attitudes" that are based on an emotional response function of the stimulus ("attitude object") that may be established through automatic or indirect conditioning processes rather than conscious verbal behaviour.

Attitudes, conceived as affective responses to words and other stimuli, may be the very sort of attitudes that "implicit tests" are best positioned to detect. The Relational Elaboration Coherence model (Barnes-Holmes et al., 2010) suggests that when participants are placed under time pressure to respond to a stimulus (as in most implicit test procedures), the most probable response will be emitted first, because the participants do not have the opportunity to engage in complex and elaborate relational responding. The distinction between “brief immediate responses” emitted under time pressure and more extended relational responding resembles the distinction between “affect based” attitudes and “cognitive based” attitudes. As such, attitudes that are primarily based on the emotional function of the stimulus will be the kinds of attitudes that are detected most directly and unambiguously by implicit measures (i.e., which enforce or instruct brief immediate responses).

While Experiment 1 examined the ability of the FAST to detect trained stimulus relations for which no affective or emotional functions had been established, Experiment 2 will consider whether or not the current FAST preparation can be used to detect functional relations between stimuli (e.g., based on shared emotional response functions). As in the previous experiment, we will examine the FAST’s ability to detect very simple laboratory established relations in order to strengthen the proof of concept established by Experiment 1. Participants will complete a simple respondent conditioning procedure to establish two emotional functions for two nonsense syllable stimuli. After a brief testing block to ensure that the appropriate conditioning has taken place, participants will then complete three pairs of FAST test blocks (i.e. consistent/inconsistent blocks) and four baseline phases – one before the first FAST test blocks, and one after each pair of test blocks.

Method

Participants

21 volunteers participated in the study as participants. Eight were eliminated due to failing to reach performance criteria during the FAST test blocks (see Results). Of the remaining participants, two further participants were eliminated due to failing to reach criterion in the Function Testing Phase. Of the 11 remaining volunteers, eight were male and five were female. Ages ranged from 18 to 25 years (M = 21.55, SD = 2.697). The participants were recruited both from acquaintances of the experimenter and using a ‘snowballing’ technique whereby participants were asked to recruit a further participant for participation. Informed consent was obtained in writing from all participants.

Apparatus and Stimuli

All phases of the experiment were presented to participants on an Apple Macbook laptop using the Psyscope software package (see Experiment 1). Stimuli consisted of both photographic images and nonsense syllables. Twenty-two nonsense syllables were utilized (see Appendix 2) arbitrarily assigned to their roles as samples, comparisons, and FAST test stimuli (see below). These will be referred to hereafter using alphanumerics. Twelve images from the International Affective Pictures System (IAPS; Lang, Bradley, & Cuthbert, 1999) were also used as unconditioned stimuli (US). Six erotic images were classified in the IAPS under the headings; romance, erotic couple, and couple. Six aversive images were classified in the IAPS under the headings; roaches, attack dog, disabled, electric chair, distressed femme, and attack.

Ethics

All participants were presented with and signed a consent form before proceeding to first phase of the experiment (See Appendix 1). The consent form specified that some of the images that would be used in the study would be of an aversive or erotic nature. Participants were told informally that performance on the task would not allow the researcher to make any individual psychological assessments but may allow for group patterns to be identified. All participants were informed of the true nature of the study after participation and were given the opportunity to withdraw their data at that stage.

General Experimental Procedure

The experiment consisted of three phases. Phase 1 (Function Training) consisted of a respondent conditioning procedure to establish emotional functions (erotic and aversive) for two nonsense syllables (i.e., unconditioned stimuli). Phase 2 (Function Testing) determined if the desired stimulus functions had been successfully established. The third and final phase involved the administering of three consecutive FAST procedures. This consisted of three pairs of test blocks (one consistent block and one inconsistent block), and four baseline blocks – one presented before each pair of test blocks, and one after each pair of test blocks.

Phase 1: Function Training

In this phase, participants were exposed to a function training protocol to establish erotic functions for the nonsense syllable “CUG” (A1) and aversive functions for the nonsense syllable “JOM” (A2). At the onset of this phase, the following instructions were presented:

In a moment, some words and images will appear on this screen. Your task is to look at these items carefully and remember what you see. It is very important that you continue to watch the screen at all times. The screen will be blank for various periods of time. This is normal.

After each picture has been presented, you will be required to press the space bar on the computer keyboard to continue. Please make sure you know where the spacebar is before you begin.

Remember – It is very important that you pay close attention to what is happening on the computer screen. If you have any questions, please ask them now. When you are ready to begin, please click the mouse button.

Function training consisted of a block of 32 trials presented in the form of respondent conditioning trials. Each trial began with the presentation of either A1 or A2 in the centre of the screen in 48pt red (A1) or blue (A2) font on a black background. The stimulus remained on screen for 3s before being replaced by a black screen for 3s. The stimuli were presented in a quasi-random sequence such that no stimulus appeared more that twice in succession.

Next, an image stimulus was presented in the centre of the screen. The presentation of A1 was followed by the presentation of an erotic stimulus. The presentation of A2 was followed by the presentation of an aversive stimulus. While image stimuli were being presented, the corresponding nonsense syllable was simultaneously displayed in the top right corner of the screen. Thus, the functioning training procedure consisted of a combination of trace and simultaneous conditioning, in order to maximise conditioning effects. The stimulus remained on screen for 3s. The erotic or aversive images were presented quasi-randomly such that no image was displayed twice in succession following the relevant stimulus.

When the image had disappeared from the screen the command “please press the space bar” appeared in the centre of the screen and remained on screen until the space bar was pressed. The next trial commenced immediately after the participant produced the required observation response.

Phase 2: Function Testing

The Function Testing Phase probed for the formation of the functional stimulus classes predicted by Phase 1 (i.e., A1-erotic image, A2-aversive image) using a matching-to-sample procedure.

The following instructions were presented to participants at the onset of Phase 2:

In a moment some words and images will appear on screen. Your task is to look at the word at the top of the screen and chose one of the images at the bottom of the screen by "clicking on it" using the computer mouse and cursor. You will not be provided with any feedback at this stage. If you have any questions, please ask them now.

The function testing consisted of ten trials. In each trial, the participant was presented with either a romantic or aversive image, displayed in the upper centre of the screen. The images used in this phase were different to those utilized in the training phase. Erotic and aversive images were selected quasi-randomly by the computer software such that no image appeared more than twice in succession. Two target stimuli, A1 and A2, were displayed in the bottom left and right of the screen as comparison stimuli. Comparison stimulus positions were randomly alternated across trials. Participants were required to click on either A1 or A2 to proceed to the next trial. The criterion for passing this phase was to produce 16 correct responses out of 20 (80%).

Phase 3: The Function Acquisition Speed Test

Phase 3 consisted of three consecutive Function Acquisition Speed Tests. The first of these was the critical test. The subsequent tests were administered to consider the robustness of FAST effects across repeated immediate exposures. The basic FAST presentation consists of a baseline block, two test blocks (consistent and inconsistent) and a further baseline block. In this experiment, participants were exposed to three sets of test blocks, each preceded by a baseline block, with a final baseline block at the end. (Thus producing the sequence; Baseline

1, Test Blocks 1, Baseline 2, Test Blocks 2, Baseline 3, Test Blocks 3, Baseline 4- See Figure 3 below).

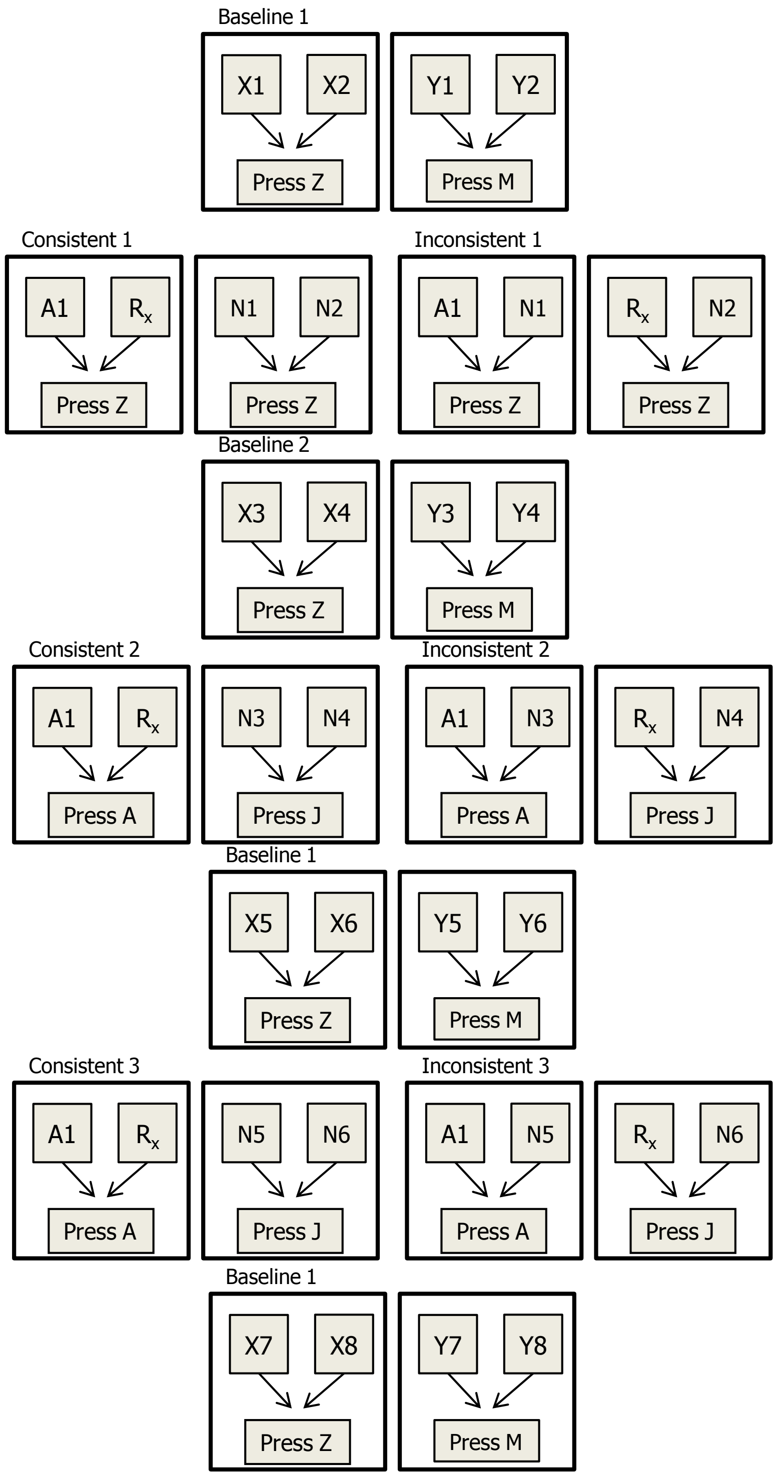


Figure 3. A schematic of the block order and various response functions assigned to stimuli during baseline and FAST blocks. "R_x" represents a randomly selected romantic image. The panels, read from top to bottom, represent each FAST test presented in sequence.

General FAST Block procedure.

Feedback, trial order, response windows, and maximum trial limit were identical to the *General FAST Block Procedure* used in Experiment 1. The presentation of nonsense word stimuli was identical to that used in Experiment 1. Image stimuli (used in the test blocks) were presented in the centre of the screen.

Baseline Blocks:

The baseline blocks each involved novel and unique nonsense syllable stimuli with which the participants had no previous experience. Baseline 1 employed X1, X2, Y1, Y2 as stimuli, while Baseline 2 employed X3, X4, Y3, Y4 as stimuli, and so on for Baselines 3 and 4. Baseline block 1 required participants to learn to "press left" when presented with X1 or Y1, and to "press right" for X2 and Y2. Baseline block 2 required common functions for X3/Y3 and X4/Y4, and so on for subsequent baseline blocks.

Four baseline blocks were presented. The repeated administration of FAST tests with baseline blocks allows for the assessment of the stability of baseline rates of function acquisition across time. Administering four baseline phases also has the advantage of allowing for the calculation of a mean baseline acquisition rate if these proved to be unstable across exposures. The instructions presented at the onset of each baseline phase where identical to those used in Experiment 1.

Test Blocks

The test blocks utilized the A1 stimulus, six of the erotic images from Phases 1 and 2, and six further novel nonsense syllables (N1/N2, N3/N4, and N5/N6). One of these blocks (consistent) established two functional response classes (A1-Erotic Image and N1-N2) that were consistent with the relations predicted given Phase 1. The other block established two functional response classes that were inconsistent with the relations trained in Phase 1 (A1-N1 and Erotic Image-N2). The order of the consistent and inconsistent blocks was randomized across participants. On the second and third FAST runs, the novel stimuli were changed (N3/N4 for FAST 2, N5/N6 for FAST 3).

The following instructions were presented to participants at the onset of each FAST block.

In the following section, your task is to learn which button to press when a word or a picture appears on screen. IMPORTANT: During this phase you should press only the A key or the J 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.

Results

Function Testing

Two participants (S10 and S16) failed to reach the criterion of 80% correct responses in the function testing phase (Phase 2). These participants were eliminated from further data analysis, as the number of participants who failed at this phase was too small to serve as a meaningful control group.

Baseline Blocks

Table 1 shows the number of trials required for each participant to reach the fluency criterion. These data illustrate higher baseline scores than were evident in Experiment 1. A related samples Friedman's 2-Way Analysis of Variance by Rank showed that the distributions of baseline scores between the four runs were not significantly different ($z=.727$, $p= .987$), indicating that there were no significant practice effects. A large number of participants failed to complete one or more baseline phases in fewer than the programmed limit of 100 trials. This resulted in surprisingly large mean baseline scores for each participant, which would have led to distorted SoR scores. This is possibly due to the fact that the Baseline phases utilised only nonsense syllable stimuli, while the FAST test blocks used a combination of the two. As such, the Baseline Blocks may not have been valid in their usual role as a baseline index of how well a participant performs with neutral stimuli. As such, the analysis for this experiment will focus on the difference scores as the primary data.

Table 1: Trials to criterion and means for each participant on each of the Baseline blocks.

Sub	Baseline 1	Baseline 2	Baseline 3	Baseline 4	Mean
1	100	32	13	100	61.25
5	15	100	18	100	58.25
6	22	100	48	69	59.75
9	53	100	100	100	88.25
12	34	72	100	59	66.25
14	37	21	100	100	64.50
15	100	13	19	31	40.75
18	95	18	39	20	43
20	28	32	100	13	43.25
21	35	58	37	32	50.5
Mean	51.9	54.6	57.4	62.4	57.585

FAST test blocks

Each participant completed 3 pairs of FAST blocks (one consistent and one inconsistent per pair). Table 2 shows the trials to criterion for each participant on each block of the FAST, and the resulting difference scores. Difference scores were calculated by subtracting trials to criterion on a consistent block from the trials to criterion on the corresponding inconsistent block.

FAST Test Blocks 1. On the first test run, three of ten participants showed a difference score in the expected direction. The mean difference score on this FAST run was -12.111 (SD = 39.419). A one sample Wilcoxon signed ranks test indicated that the difference scores were not significantly different from zero (z = 8.434, p = .260). The mean difference in the expected difference was 48.333, while the mean difference in the unexpected direction was -22.5.

*Table 2: Trials to criterion, difference scores (Incon – Con), and means for each participant on the FAST test blocks. A negative difference score represents a difference in the unexpected direction.*Data missing due to software error.*

Participant	FAST 1			FAST 2			FAST 3			Means		
	Con	Incon	Diff	Con	Incon	Diff	Con	Incon	Diff	Con	Incon	Diff
1	79	29	-50	14	29	15	27	71	44	40	43	3
5	18	69	-51	20	99	79	13	53	40	17	73.6	22.6
6	93	23	70	18	49	31	59	22	-37	56.6	31.3	21.3
9	39	20	-19	14	41	27	52	44	-8	35	35	0
12	50	87	37	11	83	72	65	79	14	42	83	41
14	13	12	-1	34	38	4	14	16	2	20.3	22	1.6
15	71	*	*	28	20	-8	87	45	-42	62	32.5	-25
18	50	37	-13	31	51	20	12	35	23	31	41	10
20	23	22	-1	23	35	12	64	60	-4	36.6	39	2.3
21	78	35	38	18	12	-6	17	28	9	37.6	25	13.6
Mean	51.4	37.11	1.11	21.1	45.7	24.6	41	45.3	4.1	37.8	42.7	9.9

FAST Test Blocks 2. On the second test run, eight out of ten participants showed a positive difference score (i.e. a difference score in the expected direction). The mean difference score on this run was 24.6 (SD = 29.718). A one sample Wilcoxon signed ranks test showed that the differences scores were significantly different from zero (z = 2.293, p=.022). The mean difference in the expected direction was 32.5, while the mean difference in the unexpected direction was -7.

FAST Test Blocks 3. On the third and final FAST run, six out of ten participants showed a difference score in the expected direction. The mean difference score on this run was 4.3 (SD = 28.701). A One sample Wilcoxon signed ranks test indication that the difference scores were not significantly different from zero on a group level (z=.663, p=.508). The mean difference in the expected direction was 22. The mean difference in the opposite direction was -22.75.

Summary

Over the course of three exposures to the FAST test, an unstable pattern of test performances were observed. The first and third runs showed no statistically significant effects, with the only significant FAST effect emerging on the second FAST run. Further, in the first FAST run, the majority of participants showed a difference score in an unexpected direction, while in the third, half of the participants took longer to complete the inconsistent block. Overall, this suggests problems with experimental control in the current experiment.

Discussion

The current experiment aimed to demonstrate that the FAST methodology was capable of detecting relations between stimuli which shared an experimentally trained function. While the FAST showed the expected effect on its second run, it failed to show any significant effect in either direction on the first and third runs. Given the sensitivity of the procedure in the first experiment, this is a surprising result.

The most likely reason for the failure of the procedure to produce the expected effect is that the functional relations were not sufficiently trained in the first phase. However, each participant successfully completed the testing phase, and was able to identify the functions attached to each nonsense stimulus as trained. If inadequate training is to be implicated in the failure to generate a FAST effect, this would suggests that the FAST procedure is not particularly sensitive to newly formed relations. To some extent, this is not surprising. The FAST procedure presents two sets of reinforcement contingencies in the test blocks. The consistent block contingencies continue to reinforce the same functional response for related stimuli. However, the inconsistent block reinforces different responses. As such, newly trained relations would be most vulnerable to the disruption presented by the inconsistent block. In effect, participants might be expected to learn quickly under the inconsistent reinforcement contingencies, thus producing little or no FAST effect. However, this same point should apply to Experiment 1, but in that study the FAST was successful in detecting the trained relations. As such, it is necessary to look deeper to identify possible sources for the loss in stimulus control observed during the current experimental preparation.

One clue may lie in the fact that the FAST test blocks in Experiment 2 utilised a mixture of image stimuli and nonsense syllable stimuli. In the consistent block, for example, participants were required to respond with a left key press on presentation of an erotic image or “CUG” (consistent with their training in Phase 1) while responding with a right key press to on presentation of the other two (novel) nonsense syllables “MIP” and “KON.” While this was consistent with their initial training, the fact that three of the stimuli were three-letter nonsense syllables while one was an image means that the participants could have been responding to the natural topographical similarity between the nonsense syllables. In other words, these may have formed a natural stimulus class based on formal features. The similarity of the stimuli in itself may have functioned as a behavioural disrupter, and thus provided a confounding source of control. If this analysis is correct, it most certainly could account for the current outcome. The FAST procedure relies on the fact that the differing reinforcement contingencies between the two test blocks are the only source of behavioural disruption. If variance in test scores is caused by other factors, then the result will be masked by behavioural noise (i.e., variability). While it may be impossible to eliminate all sources of variance, one as powerful and obvious as stimulus similarity must be controlled for in future experimental design.

As noted in the Results, the high number of trials required to complete baseline blocks made the calculation of a meaningful SoR index impossible. A baseline phase which is a true and accurate assessment of the participant’s ability to complete the task *as presented to them in the test blocks* is a core part of the FAST methodology. Without such a baseline, it is difficult, if not impossible, to judge a participant’s scores meaningfully because the raw difference in learning rates across the main FAST blocks, must be adjusted to correct for individual differences in intellectual skills or manual dexterity.

The findings of this experiment also seem to suggest that the pre-programmed limit of 100 trials may not be sufficient for all participants to complete a test block. Nearly half of the participants failed to complete test blocks in 100 trials or less. This may be as a result of the issues outlined above, or due to problems with participant attention and motivation in what was a longer procedure than Experiment 1.

Due to these failures of experimental design, it is difficult to draw any strong conclusions from the current findings. However, this experiment did serve to illustrate the sensitivity of the procedure to issues of stimulus selection, and made it clear that baseline blocks needed to mirror the structural properties of main test blocks.

Experiment 3

Using the FAST to detect derived relations established in the laboratory

Experiment 1 showed that directly established relations between stimuli could be identified using the FAST procedure. However, not all relations of interest to psychologists might be directly established by the verbal community. Some relations may have been derived by the participant because they have been merely implied by the verbal community in the absence of any direct social reinforcement. As an example, the parents of a child may have regularly referred to people of *Irish* origin as *drunkards* while speaking to their child. In other contexts, the parents, or other individuals, may have referred regularly to *drunkards* as *ignorant.* These contingencies parallel a linear (A-B-C) stimulus matching sequence in which a derived relation between the terms *Irish* and *ignorant* might be expected to emerge. While there may have been no occasion on which an individual was instructed that the Irish are Ignorant, this implied or derived relation may nevertheless be detectable using a FAST procedure.

The use of derived equivalence relations as a laboratory analogue of verbal relations of interest to social researchers (e.g., in the context of attitude research) is supported by a growing body of research which suggests that derived relations function in the same way as semantic relations in the vernacular and share the same functional properties. For instance, research using event related potentials (ERPs) as a dependent measure of equivalence class formation has shown that the neural correlates of deriving relations and semantic processing are similar (Barnes-Holmes et al., 2005; Haimson, Wilkinson, Rosenquist, Ouimet, & McIlvane, 2009). Similar findings have been made in relation to fMRI measures of stimulus equivalence class formation (Dickins et al., 2001). Several studies have also shown the emergence of derived relational responding repertoires to be practically synonymous with the emergence of natural language in humans (Barnes, et al., 1990; Devany, et al., 1986; Hayes, 1989; Lipkens et al., 1988). Other research has found it difficult to demonstrate stimulus equivalence in animal populations (e.g. Dugdale & Lowe, 2000; Lionello-DeNolf & Urcuioli, 2002). The concept of derived relations has, therefore, been used by several behaviour analysts to develop models of meaning (Bortoloti & de Rose, 2009) as well as to understand grammar and syntax from a behavioural perspective (e.g., Barnes-Holmes, Barnes-Holmes, & Cullinan, 2000; Barnes & Hampson, 1993; Hayes et al., 2001). Most recently, Bortoloti and de Rose (2012) used an implicit test procedure (the IRAP) to confirm that stimuli participating in the same equivalence relation were semantically related. Thus, the concept of the derived equivalence relations may serve as an appropriate laboratory analog of implicit verbal relations as conceived and assessed by social-cognitive researchers.

The current study was designed to examine the sensitivity of the FAST procedure in identifying the existence of a derived stimulus equivalence relation between two nonsense syllable stimuli related indirectly to each other following exposure to a One-to-Many stimulus equivalence training and testing procedure. Given the issues that arose in Experiment 2, when employing pre-established, familiar and pictorial stimuli in FAST procedures, this Experiment was also intended to return the focus of research to laboratory created arbitrary relations between unfamiliar nonsense syllable stimuli until some of those issues are ironed out in more basic research.

Method

Participants

Twenty four participants were recruited from the undergraduate population at NUI Maynooth as participants. Of the twenty four who began the study, seventeen of these passed equivalence training and testing, but one was eliminated due to failing to complete a FAST test block in less than 100 trials (see below). The remaining seven participants were employed as control participants, but one of these was also eliminated due to failing to complete a FAST test block in less than 100 trials. Of the remaining twenty two participants whose data where analyzed, 11were male and 11 were female. Ages ranged from 18-48 (Mean = 24.09, SD = 7.909).

Apparatus and Stimuli

All phases of the experiment were presented to participants on an Apple Macbook laptop using the Psyscope software package (see Experiment 1). Stimuli consisted of twenty-eight nonsense syllables (see Appendix 2) arbitrarily assigned to their roles as samples, comparisons, and FAST test stimuli (see below). These will be referred to hereafter using alphanumerics.

Ethics

All participants were presented with and signed a consent form before proceeding to first phase of the experiment (See Appendix 1). Participants were told informally that performance on the task would not allow the researcher to make any individual psychological assessments but may allow for group patterns to be identified. All participants were informed of the true nature of the study after participation and were given the opportunity to withdraw their data at that stage.

General Experimental Procedure

The experiment consisted of three phases. Phase one (equivalence training) consisted of a matching-to-sample protocol to establish two 3-member equivalence classes, each containing three nonsense syllables (A1-B1-C1 and A2-B2-C2.) Phase two (equivalence testing) tested for the derived relations emergent from Phase 1 (i.e., A1-B1, B1-A1, A2-B2 and C2-B2). Phase three (FAST) consisted of three runs (exposures) of a pair of FAST test blocks, separated by a baseline block, and with an additional baseline block presented at the end of the entire procedure. Multiple runs of the FAST were employed so that the stability of participant performance across time could be considered.

Phase 1: Equivalence Training

In this phase, participants were exposed to a matching-to-sample procedure designed to establish two 3-member equivalence classes according to a one-to-many (sample-as-node) protocol. Each of four training trials appeared eight times in a quasi-random order, for a total of 32 trials (i.e., no more than two successive exposures to nay one task). The relations trained were; A1-B1 (B2), A2-B2 (B1), A1-C1 (C2), and A2-C2 (C1), where unreinforced choices are in parentheses, as detailed in Figure 4, below.

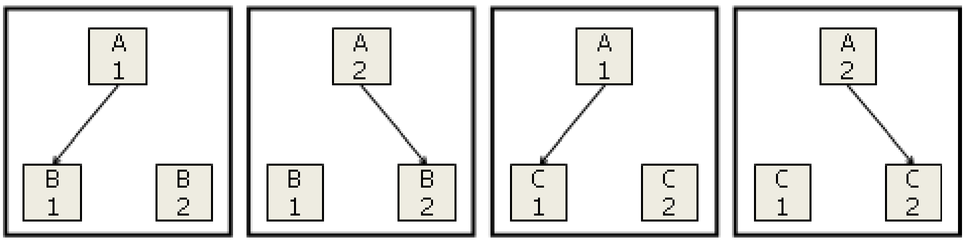


Figure 4. The stimulus relations trained during Phase 1. The arrows indicate reinforced matching responses.

Participants were presented with the following instructions at the onset of Phase 1:

In a moment some words will appear on this screen. Your task is to look at the word at the top of the screen and choose one of the two words at the bottom of the screen by "clicking on it" using the computer mouse and cursor. During this stage the *computer will provide you with feedback on your performance. You should try to get as many answers correct as possible. If you have any questions please ask them now. When you are ready please click the mouse button.*

Trial presentation was in the same format as in Experiment 1, with the exceptions noted above (i.e. one-to-many experimental arrangement) Participants were required to complete a block with 30/32 correct responses. If participants failed to meet criterion, the training block was repeated until criterion was reached.

Phase 2: Equivalence Testing

During equivalence testing, probes for the unreinforced formation of B-C relations and C-B relations (combined symmetry and transitivity) were presented (see Figure 5, below). The testing format was similar to the training procedure. However, no corrective feedback was provided. Again, participants were required to reach a criterion of 30/32 correct responses. If participants failed to reach criterion within 4 testing blocks, they were classified as control participants and they proceeded to Phase 3 as normal.

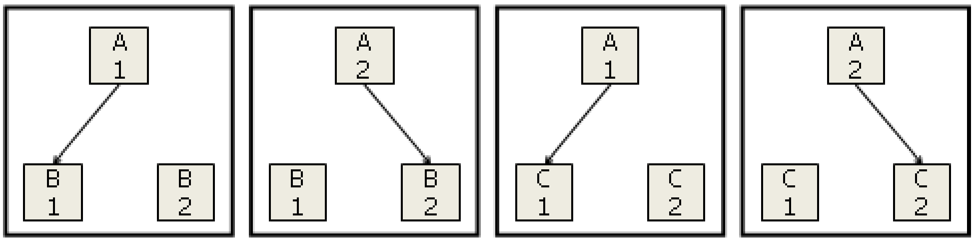


Figure 5. The derived stimulus relations probed for during equivalence testing. Arrows indicated correct choices.

Phase 3: The Function Acquisition Speed Test

Phase 3 consisted of three consecutive Function Acquisition Speed Tests. The first of these was the critical test. The subsequent tests were administered to consider the robustness of FAST effects across repeated immediate exposures. The basic FAST presentation consists of a baseline block, two test blocks (consistent and inconsistent) and a further baseline block. In this experiment, participants were exposed to three sets of test blocks, each preceded by a baseline block, with a final baseline block at the end, thus producing the sequence; Baseline 1, Test Blocks 1, Baseline 2, Test Blocks 2, Baseline 3, Test Blocks 3, Baseline 4 (see figure 6, below)

General FAST Block Procedure

Stimulus presentation, feedback, trial order, response windows, and maximum trial limit were identical to the *General FAST Block Procedure* used in Experiment 1. However, the criterion for completion of a block was altered. Participants were exposed to trials until a criterion of 9/10 correct in a row was reached. That is to say, participants were required to produce correct responses across any contiguous sequence of 10 trials, with no more than one error in that 10-trial sequence. If the participant reached this criterion, then the block ended automatically and the instructions page for the next block would be presented. Figure 6 below detailed the block order and required response functions for the FAST procedure.

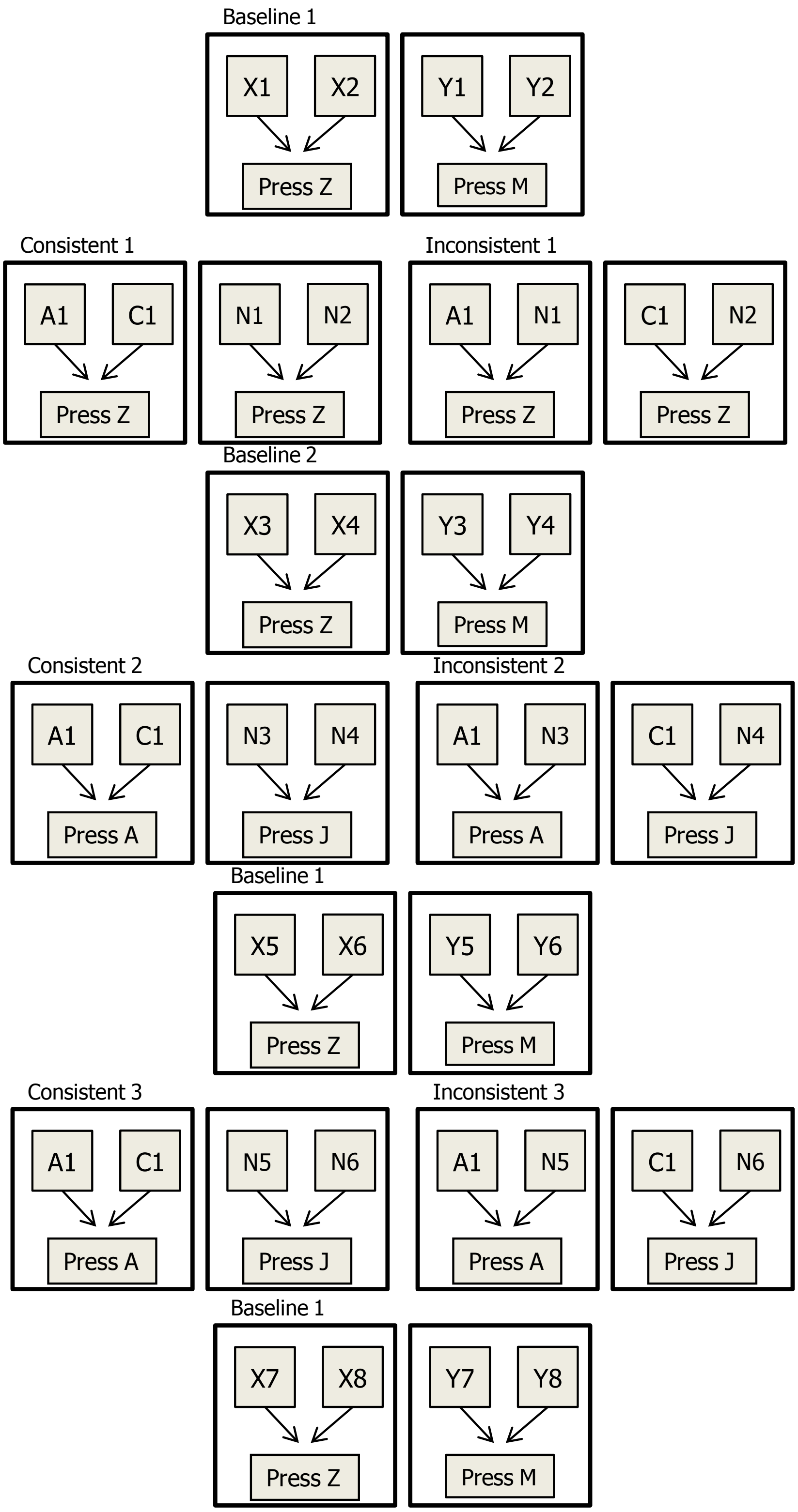


Figure 6. A schematic of the block order and various response functions assigned to stimuli during baseline and FAST blocks. The panels, read from top to bottom, represent each FAST test presented in sequence.

Baseline Blocks

The procedure used for the Baseline Blocks was identical to that used in Experiment 2.

Test Blocks

The test blocks utilized the A1 and B1 stimuli from the Equivalence Training and Testing Phases, and two further novel nonsense syllables, N1 and N2. One of these blocks (consistent) established two functional response classes (A1-B1 and N1-N2) that were consistent with the derived relations predicted given Phase 1. The other block established two functional response classes which were inconsistent with the relations trained in Phase 1 (A1-N1 and B1-N2). The order of the consistent and inconsistent blocks was randomized across participants.

The stimulus presentation procedure used in the test blocks was identical to that used in the Baseline Blocks – only the stimuli used differed between blocks. The test blocks also utilized different response keys (*z* and m) to prevent any conflicting response histories across baseline and test blocks.

The instructions presented to participants at the onset of the FAST Baseline Blocks and Test Blocks were identical to those presented in Experiment 1.

To summarize, participants were exposed to a series of blocks in the following sequence; Baseline 1, FAST 1, Baseline 2, FAST 2, Baseline 3, FAST 3, and Baseline 4. After the completion of each block, the instructions page for the next block appeared, allowing the participant to begin the next block whenever they were ready without interruption from the experimenter. Upon completion of the final baseline block, a page appeared thanking the participant for their participation and instructing them to contact the experimenter.

Results

Data for twenty-two participants was analysed (see Participants section). Participants required a mean of 6.35 blocks to complete equivalence training. The 16 participants who successfully passed equivalence testing required a mean of 1.38 testing blocks to pass. The remainder all continued to fail on equivalence testing blocks on the fourth and final exposure. The number of training and testing blocks required by each participant is summarized in Table 1. Data for participants who successfully passed equivalence training (i.e., experimental Participants 1-16) will be analysed first below.

Table 1: Number of blocks of equivalence training and testing blocks required to reach criterion by each participant. F indicates that the performance resulted in a fail classification

Sub	1	2	3	4	5	6	7	8	9	10	11	12
Training	6	4	3	2	3	6	2	3	6	3	5	3
Testing	1	1	1	1	1	1	1	2	2	2	1	1
Sub	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**		
Training	6	1	5	3	3	2	8	7	11	6		
Testing	2	1	1	2	4 (F)	4 (F)	4 (F)	4 (F)	4 (F)	4 (F)		

Experimental Participants (i.e., passed equivalence testing)

Participants completed four baseline blocks during the FAST procedure. Table 2 shows the number of trials required by each experimental participant to reach the fluency criterion (10 in a row correct, with one error allowed) on each exposure to a baseline block.

Table 2: Number of trials to criterion on FAST baseline blocks for experimental participants.

Participant	Run 1	Run 2	Run 3	Run 4	Mean
1	41	42	10	100	48.25
2	26	24	31	34	28.75
3	26	27	10	10	18.25
4	29	17	19	20	21.25
5	12	31	12	14	17.25
6	25	21	11	24	20.25
7	16	16	10	19	15.25
8	40	48	10	23	30.25
9	92	60	88	67	76.75
10	81	10	10	52	38.25
11	100	26	13	17	39
12	100	63	35	100	74.5
13	82	27	15	45	42.25
14	12	24	27	16	19.75
15	11	19	15	16	15.25
16	11	43	26	24	26
Mean (SD)	44	31.13	21.38	36.31	33.2
	(34.31)	(15.66)	(19.61)	(29.22)	(19.43)

These data were analysed using a Repeated Measures ANOVA to test for differences in baseline block performance across time. There was a main effect for time ($F=3.759$, $p= 0.017$), with significant differences between Baselines 1 and 3 ($p= 0.014$) and 2 and 3 ($p= 0.03$). While there was an increase in trial requirements from Baseline 3 to Baseline 4, this was not a statistically significant difference. In effect, baseline trial requirements varied to some extent across runs, although the general trend was towards lower trial requirements (i.e., practice effects). A mean baseline trial requirement score was calculated for each participant and employed in calculating the Strength of Relation indices for each FAST run (see below).

FAST Run 1. Table 3 outlines the trial requirements for each participant on each block of the three FASTs administered. Eleven of the sixteen participants showed a faster rate of response function acquisition on the consistent compared to the inconsistent block, as expected. The average trial requirement difference in the positive direction was 24. By contrast, the difference in the unexpected direction was smaller (Mean = -14). A Wilcoxon signed ranks test was conducted, showing that the acquisition rate differential was significant at the group level ($z = 1.707$, $p = 0.04$, one tailed) and in the expected direction.

FAST Run 2. Ten of the sixteen participants showed the expected response rate differential between the consistent and inconsistent blocks. One participant showed no difference, while the remaining five participants showed faster rates of acquisition during the inconsistent blocks. The mean trial requirement difference in the expected direction was 16.67, while the mean difference in the unexpected direction was -13.4. A Wilcoxon signed ranks test showed that this difference was not significant at the group level ($z = -1.364$, $p = 0.08$, one tailed).

FAST Run 3. Eleven of the sixteen participants showed a faster rate of response function acquisition on the consistent compared to the inconsistent block, as expected. The average difference in the positive direction was 14.4. By contrast, the mean difference in the unexpected direction was smaller (Mean = -6.6). A Wilcoxon signed ranks test was conducted, showing that the acquisition rate differential was significant at the group level ($z = -1.907$, $p = 0.025$, one tailed).

Table 3: Number of trials to criterion on FAST test blocks, difference scores, and block order (1= consistent block first) for experimental participants.

Participant	FAST 1			Order	FAST 2			Order	FAST 3			Order
	Diff	Con	Incon		Diff	Con	Incon		Diff	Con	Incon	
1	-8	39	31	1	19	34	53	1	22	22	44	2
2	17	21	38	1	8	10	18	1	-20	38	18	2
3	7	24	31	1	15	13	28	1	6	15	21	1
4	-13	31	18	1	-1	14	13	2	6	14	20	1
5	-6	16	10	1	0	11	11	1	3	12	15	1
6	2	13	15	1	3	13	16	1	-2	12	10	2
7	25	13	38	2	9	14	23	2	-5	24	19	2
8	6	18	24	1	-3	18	15	2	-1	11	10	1
9	63	16	79	1	-37	64	27	2	33	67	100	2
10	18	16	34	1	-24	34	10	2	3	13	16	1
11	-15	33	18	1	-2	15	13	1	6	11	17	1
12	57	32	89	1	76	10	86	1	-5	20	15	2
13	40	52	92	1	5	41	46	2	53	25	78	2
14	-18	41	23	2	7	11	18	1	4	10	14	2
15	4	15	19	2	8	15	23	2	22	15	37	1
16	55	17	72	2	1	21	22	1	4	16	20	1
Group Mean (SD)	14.63 (26.56)				5.25 (23.342)				8.06 (17.203)			

The Strength of Relation Index. The Strength of Relation (SoR) index developed in Experiment 1 is a simple calculation which places each participant's FAST test difference score in the context of their own baseline scores. A participant's baseline performance indicates the speed at which a functional response class is formed in the absence of any pre-experimental history involving relations between the relevant stimuli. Taken alone, raw difference scores calculated for the FAST can be misleading as this difference fails to reflect individual differences in baseline acquisition rates for tasks of this kind (e.g., a difference score of 4 is highly meaningful for a participant whose baseline acquisition rate is rapid, while the same difference score is less meaningful if baseline acquisition rates are slow).

A slightly different way of calculating the SoR index was used herein Experiment 3. This is because the SoR indices obtained in Experiment 1 were small and non-standardised. Functionally, however, the SoR formula will reflect the same properties in both experiments. Importantly, the reader should not compare SoRs directly across Experiments 1 and 3 due to this difference in how the SoRs were calculated.

The current SoRs were calculated for each participant by dividing each participant's raw difference score (inconsistent block – consistent block) by the natural logarithm of that participants' mean baseline trial requirement. This creates an index which is zero when a participant has equal acquisition rates on consistent and inconsistent blocks, is positive when inconsistent block acquisition rates are higher than consistent block acquisition rates, and negative when the opposite is the case. The SoR indices for all experimental participants for each run of the FAST can be seen in Table 4. One sample t-tests were conducted on these scores for each run of the FAST. The SoR indices for the first ($t=2.174$, $p = 0.023$, one tailed) and third ($t=1.896$, $p = 0.038$, one tailed) FAST runs were significantly different from zero, while the SoR calculated for the second run was not ($t=1.086$, $p = 0.147$, one tailed).

In summary, the FAST procedure was sensitive to the emergent non-reinforced stimulus relations at the group level on the first (i.e., critical) exposure. The FAST effect during this first exposure was significant using both raw difference scores and differences in SoR indices from zero.

Table 4: Strength of Relation indices and block order (1= consistent block first) for experimental participants. A positive index score indicates a predicted FAST effect.

Participant	FAST 1	Order	FAST 2	Order	FAST 3	Order
1	-2.06	1	4.9	1	5.68	2
2	5.06	1	2.38	1	-5.95	2
3	2.41	1	5.16	1	2.07	1
4	-4.25	1	-0.33	2	1.96	1
5	-2.11	1	0	1	1.05	1
6	0.66	1	1	1	-0.66	2
7	9.18	2	3.3	2	-1.84	2
8	1.76	1	-0.88	2	-0.29	1
9	14.51	1	-8.52	2	7.6	2
10	4.94	1	-6.59	2	0.82	1
11	-4.09	1	-0.55	1	1.64	1
12	13.22	1	17.63	1	-1.16	2
13	10.68	1	1.34	2	14.16	2
14	-6.03	2	2.35	1	1.34	2
15	1.47	2	2.94	2	8.07	1
16	16.88	2	0.31	1	1.23	1
Group Mean (SD)	3.8894 (7.156)		1.52 (5.62725)		2.231 (4.708)	

Control Participants (i.e., failed equivalence testing)

Participants 17-22 failed to pass the equivalence testing phase within the predetermined limit of four test blocks. Table 5 summarizes the number of trials to criterion required by each participant while completing the baseline blocks.

Table 5: Number of trials to criterion on FAST baseline blocks for control participants.

Participant	Baseline 1	Baseline 2	Baseline 3	Baseline 4
17	32	10	33	30
18	19	20	12	12
19	37	92	74	100
20	30	41	25	.
21	53	39	34	27
22	48	100	14	66
Group Mean (SD)	36.5 (12.438)	50.33(37.324)	32 (2.548)	47 (35.651)

FAST Run 1. Table 6 outlines the trial requirements for each participant on each block of the FAST. Four of the six participants who failed equivalence testing showed a faster rate of response function acquisition on the consistent compared to the inconsistent block (i.e., P17, 18, 19, 22) while the remaining two participants showed the reverse. Generally, difference scores were smaller than those observed for experimental participants with a mean of 7.33 (approximately half the mean difference score observed during run 1 of the FAST using experimental participants). A Wilcoxon signed ranks test was conducted, showing that the acquisition rate differential was not significant at the group level ($p = 0.673$).

FAST Run 2. On the second run of the FAST, only one participant showed a faster rate of acquisition on the consistent block (i.e., a standard FAST effect). Four participants showed a faster rate of acquisition on the inconsistent block (not predicted), and one showed no difference. Due to a computer error, data for P22 were not recorded during the second run of the FAST. The mean difference score was 1.2 (approximately four times smaller than that observed during run 2 of the FAST using experimental participants). A Wilcoxon signed ranks test showed that this response acquisition differential was not significant at a group level (p = 0.715).

FAST Run 3. On the final exposure to the FAST, 3 participants showed a faster rate of acquisition on the consistent block, while 3 others showed the opposite effect. Participant 20 opted to end his participation after 2 runs of the FAST and so no data is available for this participant. The mean difference score was 4.0 (approximately half of that observed during run 3 of the FAST using experimental participants). A Wilcoxon signed ranks test showed again that the response acquisition rate difference was not significant at a group level (p= 0.500).

Table 6: FAST block trials to criterion, difference scores, and block order (1= consistent block first) for control participants.

Participant	FAST 1			Order	FAST 2			Order	FAST 3			Order
	Diff	Con	Incon		Diff	Con	Incon		Diff	Con	Incon	
17	2	11	13	2	0	13	13	1	31	10	41	2
18	2	12	14	1	-2	12	10	1	6	10	16	1
19	12	33	45	1	-25	64	39	2	-3	13	10	1
20	-14	37	23	1	12	38	26	1	.	.	.	.
21	-12	30	18	2	45	34	79	2	17	17	34	2
22	54	12	66	1	.	.	.	1	-30	49	19	1
Group Mean (SD)	7.33 (24.841)				1.20 (26.414)				4 (22.95)			

The Strength of Relation Index. As would be expected from the difference scores detailed above, the SoR indices calculated for the seven participants who failed to pass equivalence testing were highly variable. Single sample t-tests were conducted comparing the SoR scores for each exposure to the FAST to zero. In each instance, the SoR scores were not significantly different from zero. (SoR 1; t = 659, p = 0.539, SoR

2; t = .152, p = 0.887, SoR 3; t = .584, p = 0.591). Table 7 summarizes the SoR indices for the control participants.

Table 7: Strength of Relation indices and block order (1= consistent block first) for control participants.

Participant	**SoR 1**	**Order 1**	**SoR 2**	**Order 2**	**SoR 3**	**Order 3**
17	0.61	2	0	1	9.49	2
18	0.73	1	-0.73	1	2.18	1
19	2.77	1	-5.78	2	-0.69	1
20	-4.04	1	-3.46	1	.	.
21	-3.29	2	12.35	2	4.67	2
22	13.36	1	.	.	-7.42	1
Group Mean (SD)	1.689 (6.277)		.476 (7.022)		1 .643 (6.29)	

Summary

Both descriptive and inferential statistical analyses of the FAST block trial requirements suggest that significantly faster acquisition of response functions was observed on the consistent block compared to the inconsistent block, but only for experimental participants who had previously derived equivalence relations between the critical stimuli. In effect, the FAST procedure would appear to be sensitive to derived relations.

Discussion

The current study expands upon the results of previous research by demonstrating that the FAST is capable of detecting derived relations between stimuli as well as directly trained stimulus-stimulus relations and increases confidence in the basic procedure. In other words, these findings support the concerns raised in Experiment 2 that natural topographical categories obtaining between the N1 and N2 stimuli can interfere with the FAST effect. Furthermore, baseline blocks that do not accurately mirror the main FAST block structures, may also lead to variable FAST outcomes. Most importantly, however, these findings suggest that derived as well as directly trained relations can be measured by the FAST method. The FAST procedure was shown to be sensitive to derived relations that had never been reinforced in the history of the participants. This is particularly exciting because derived relational responding is thought to be a fundamental aspect of human verbal behaviour (see Hayes, et al., 2001), and as such the current findings bode well for the FAST's future extension as a test for complex verbal and social histories.

In the current study, eleven of sixteen participants who passed the equivalence test showed effects in the predicted direction on their first exposure to the FAST. Those effects which were in the predicted direction were larger on average than those observed in the opposite direction. In contrast, participants who failed equivalence testing did not show any significant FAST effects at the level of means or using inferential analyses. Overall, the difference scores and SoR scores shown by control participants tended towards zero, with considerable variation around that point. In effect, control participant data showed no convincing tendency towards either positive or negative FAST effects, suggesting that FAST performances were not under the control of a history of responding to the experimental stimuli.

Experiment 4: Using the FAST to test derived relations after extensive MTS training.

In the three experiments previous, there were several participants for whom the FAST did not show a positive Strength of Relation Index (i.e., a difference in rate of acquisition in the expected direction between consistent and inconsistent blocks). In Experiment 2, it was suggested that perhaps that this could be due to poor stimulus control in the relational training, resulting in relations not being sufficiently robust. While stimulus control was improved in Experiment 3, compared to Experiment 2, it is still the case that 8 of 22 participants in Experiment 3 failed to form equivalence relations during testing. This figure is somewhat surprising given the simple nature of the equivalence classes being trained. This suggests the possibility that the equivalence training procedure was not sufficiently robust to produce reliable and enduring equivalence responding, which in turn would lead to stable and reliable FAST effects.

The previous experiment utilized a One-To-Many training structure (i.e., A-B, A-C). Although the literature is inconclusive, both One-To-Many and Many-To-One (e.g., B-A, C-A) training structures produce more positive outcomes than a linear series training structure (Arntzen, Grondahl, & Eilifsen, 2010). Nevertheless, there are a number of improvements suggested by the literature that could strengthen the equivalence training phase as employed here.

Firstly, the previous experiment did not require an observation response to the sample, a procedure that has been found to improve acquisition of discriminations in matching-to-sample training (Arntzen, Braaten, Lian & Eilifsen, 2011). Secondly, increasing the delay between sample and comparison presentation across trials has been shown to increase the yield in equivalence responding (Arntzen, 2006). Thirdly, the introduction of a fading of consequences across training could also be implemented, in order to control for extinction (see Artnzen et al., 2010).

The current experiment seeks to incorporate these elements into the equivalence training procedure in order to maximise the probability of stimulus equivalence performances emerging prior to the FAST. Further, participants will be exposed to a much greater number of trials over an extended period of time to ensure that the laboratory trained relations are highly robust. If a large number of participants fail to produce the expected FAST effect after such extensive training, it would point more conclusively to stimulus control flaws with the FAST. On the other hand, if more reliable and stronger FAST effects emerge after such an extended and improved training history, it will corroborate the suspicion that part of the variability observed in the FAST outcomes is due to instability in the underlying relations themselves.

Four other features were added to this FAST procedure in this experiment in an attempt to stabilise FAST effects. The first change relates to the duration of the response window. In the current experiment, the response window was reduced to 2000ms.The REC model (Barnes-Holmes et al., 2010) suggests that performances on implicit tests result from the fact that time constraints on participants' responses confine them to their "brief, immediate response." The more time the participant has available to respond, the more likely they are capable of engaging in further verbal behaviours to elaborate upon, and modify that response, and bring it into line with social norms or expected behaviour. As such, the IRAP has reduced the response window to 2000ms in recent years (Barnes-Holmes et al., 2010) and found this increased the stability and reliability of their results.

The second change related to the display of learning rates to participants during the task blocks. That is, during the current FAST blocks, the number of successive correct responses made was continuously displayed at the bottom of the screen. This change was included due to the surprising number of participants who had failed to complete a FAST block within 100 trials. Many of these had informally reported that they became confused because they had no way of knowing how well or poorly they were performing at any stage. In some cases this lead to high rates of non-responding (which resets one's score to zero).

The third change to the FAST procedure relates to the use of practice blocks, designed to familiarise participants with the procedure before baseline and main FAST blocks are administered. Exposure to a single practice block involving novel stimuli may help to stabilise behaviour during subsequent blocks. The use of such blocks is common practice in the IAT and IRAP. The IAT typically uses two practice blocks, in which participants' sort category items alone, or attribute items alone (Lane, Banaji, Nosek & Greenwald, 2007). The IRAP utilizes a minimum of two practice blocks, in which participants must reach a performance criterion before proceeding to the test blocks. Participants must complete a practice block with 80% correct responses and an average response time of less than 2000ms. For participants who have difficulty reaching this criterion, there may be as many as six practice blocks.

The fourth change in the current method involves the fixing of the block order (consistent first, then inconsistent). Prior experiments did not display order effects. However, Experiment 2 demonstrated the complex issues that can arise as a result of order effects. Large group level experiments such as those typically used in IAT research typically randomise block order, with the intention of washing out any noise generated by order effects across a large sample. However, the preferred approach from a behavioural perspective is to control the effects without recourse to statistical methods. The processes underlying the FAST procedure are well understood, and as such, order effects can be explained and controlled for. The inconsistent block of the FAST is designed to be a *behavioural disrupter*. The FAST method itself rests on the notion that measuring the strength of relations rests on the measurement of the behavioural momentum before and after the presentation of a behavioural disrupter. As such, it is more appropriate to fix the block order as follows; consistent first, followed by inconsistent, and acknowledge that this arrangement of procedures may result in a slight strengthening of the socially congruent relations before the disrupter is applied.

Method

Participants

Seven participants participated in this study. Two participants (S3 and S6) dropped out before completing the full procedure. Of the remaining five participants, three participants were female, and two were male. Ages ranged from 21-53 (Mean = 29.8, SD = 13.42). All participants were acquaintances of the experimenter.

Apparatus and Stimuli

All phases of the experiment were presented to participants on an *Apple Macbook*® laptop computer with a 13-inch monitor (1024 X 768 pixel resolution). Stimulus presentations were controlled by the software package *Psyscope* that also recorded all responses. Stimuli consisted of fourteen three letter nonsense syllables randomly assigned to their roles as samples, comparisons and FAST test stimuli. These will be referred to using alphanumerics in future.

Ethics

All participants were presented with and signed a consent form before proceeding to first phase of the experiment (See Appendix 1). Participants were told informally that performance on the task would not allow the researcher to make any individual psychological assessments but may allow for group patterns to be identified.

When participants consented to participate in the experiment, they were told that it would involve a series of matching tasks presented on a computer. They were told that the first several tasks would involve choosing three letter nonsense syllables by clicking on them with the mouse, and learning which stimuli went together. Participants were informed that the experiment would take place over 3 days, and that each session would last between 30 and 60 minutes, depending on how quickly they learned to respond correctly. Further, they were told that the experiment consisted of a series of stages, and that upon completion of a stage they would move up to another stage with slightly different presentation and criteria for success, but that the general form of the task would remain the same

Matching-to-Sample Training

Participants were exposed to a series of one-to-many (sample-as-node) delayed matching-to-sample training stages, each designed to form two three-member equivalence classes (A1-B1-C1) and (A2-B2-C2). There were nine stages of training. Each time a participant reached the criterion for a particular stage, they moved on to the next stage in which more stringent criteria for completion were applied. To this extent, the current procedure can be conceived of as a changing criterion design. Specifically, as stages progressed, the number of trials per blocks and the criteria for success increased (15/16, 31/32, 47/48), as did the delay between the presentation of sample and comparison stimuli (2s, 4s, 6s). In the final three stages, feedback was reduced, from 100% to 50% to 33.3%. Participants completed the nine stages over the course of three separate training sessions, over a period of 3-5 days. The training program was completed, and participants were then tested for the emergence of stimulus equivalence without feedback (See below for details). Thereafter, participants completed a FAST procedure testing for the strength of derived relations between A and C stimuli.

Prior to the onset of each phase, instructions were presented to the participants on the computer screen. The instructions described the requirements of each task, and alerted participants to changes in stimulus presentation, block length, and contingency changes as relevant. Full text of the instructions can be found in Appendix 3. Once participants had read and understood the instructions, they clicked on the mouse button to begin the first trial.

All trials were presented on a white computer screen background. At the onset of each trial, a sample stimulus appeared in the upper centre of the screen in black 24pt font. Participants were required to emit an observation response by clicking on the sample stimulus. This caused the sample stimulus to disappear, and after a delay of 2s, the comparison stimuli appeared in the bottom corners of the screen. The left/right position of comparison stimuli was randomly counterbalanced throughout the experiment. Participants selected a comparison stimulus by clicking on it, after which the screen cleared and feedback was presented for a duration of 1000ms. A correct choice was followed by the word “Correct” in 48pt green text, while an incorrect choice was followed by the word “Wrong” in 48pt red text. After the feedback had been displayed, the next trial began with the presentation of the sample stimulus. The block continued for a variable number of trials (see below) and ended when participants achieved a criterion of 1 or fewer errors across the block. If the participant did not achieve the criterion, the block was re-presented, up to a maximum of 10 cycles. If the participant did not reach criterion after 10 blocks, then they were returned to the previous stage and asked to complete it again.

The structure of the successive stages was as follows.

Stage 0: Acquisition. The acquisition stage consisted of one block of sixteen trials in which a response to the sample stimulus caused only the correct comparison to be presented on screen. Each trial type was presented four times in a quasi random order. The trial types were A1-**B1**, A1-**C1,** A2-**B2**, and A2-**C2**.

Stage 1. Blocks in this stage consisted of 16 trials. In this stage, and all stages thereafter, a response to the sample stimulus produced two stimuli in the bottom left and right corners of the screen. Participants were required to complete the block with 15 correct responses in order to complete this stage. If the participant did not achieve the criterion, the block was represented, up to a maximum of 10 cycles. If the participant did not reach criterion after 10 blocks, then they were returned to the previous stage and asked to complete it again. The four trial types were: A1-**B1**/B2, A1-**C1**/C2, A2-B1/**B2**, and A2-C1/**C2**, where emphasized stimuli represented correct choices. The tasks can be seen in Figure 7 below.

Stage 2. This was similar to Stage 1, but these blocks consisted of 32 trials and required the participant to respond correctly in 31 out of 32 trials.

Stage 3. This was similar to Stage 1, but with blocks of 48 trials, requiring 47 out of 48 correct responses.

Stage 4-6. This was similar to Stages 4-6, but the delay between the removal of the sample stimulus (after clicking on it) and the onset of the comparison stimuli was increased to 4s. Stage 4 required 15/16 correct responses, Stage 5 required 31/32 correct responses, and Stage 6 required 47/48 correct responses.

Stages 7-9. For Stages 7-9, the delay between the presentation of sample and comparisons was increased to 6s. Further, at Stage 7 and 8 only 50% of responses received feedback. Stage 7 required 15/16 correct responses to proceed. Stage 8 required 31/32 correct responses. Stage 9 required 47/48 correct responses for completion, and only 33.3% of trials received feedback.

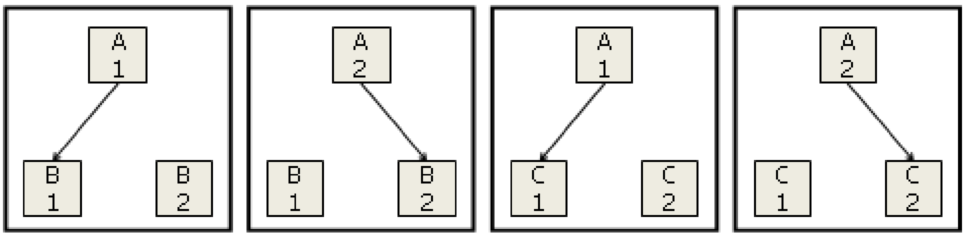


Figure 7. The stimulus relations trained during nine stages of MTS training. The arrows indicate reinforced matching responses.

Equivalence Testing

After completing the ninth stage of training, participants were tested for emergent equivalence relations. In these test blocks, all test trial types (i.e. symmetry, transitivity and equivalence trials) were intermixed in blocks of 36, with each trial type being presented 6 times in a quasi-random order. In the testing phase, no feedback was presented, and there was no delay between sample and comparison.

The trial types presented were B1-**A1**/A2, B2-A1/**A2** (Symmetry) B1-**C1**/C2, B2-C1/**C2** (Transitivity) and C1-**B1**/B2, C2-B1/**B** (Equivalence). The trial types are described in Figure 8 below.

Participants were considered to have passed if they exceeded a score of 90% correct (i.e. 33/36 or better) in a single block of testing. If a participant failed to complete equivalence testing within one block, the block was presented again, up to a maximum of four times. If a participant failed to complete the block after four successive presentations, they were classified as a control participant. (No participant required more than one presentation of this block to complete equivalence testing).

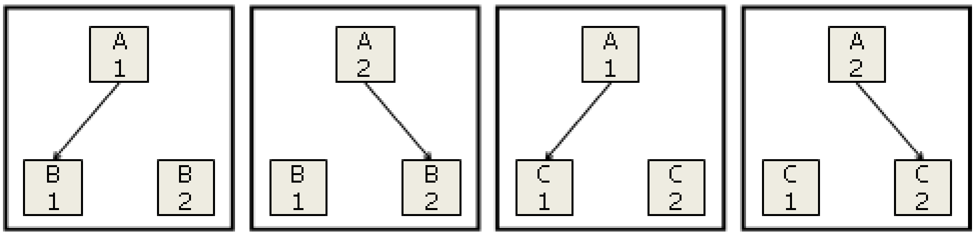


Figure 8. The derived stimulus relations probed for during equivalence testing. Arrows indicated correct choices.

Function Acquisition Speed Test

The FAST consisted a of a practice block, two test blocks and two baselines block, with one administered before the test blocks, and one afterwards. The sequence of FAST blocks is described in Figure 9 below.

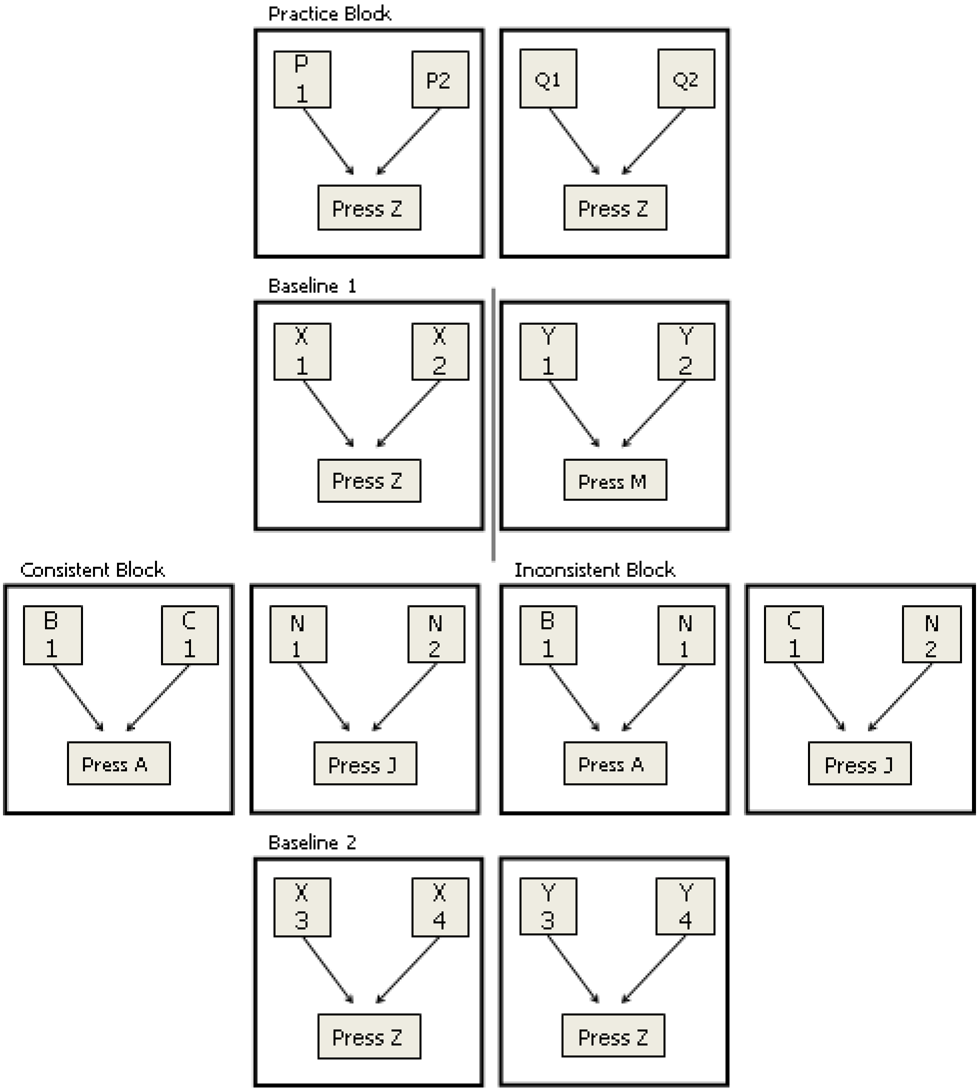


Figure 9. A schematic of the various response functions assigned to stimuli during practice, baseline and FAST blocks. The four rows, read from top to bottom, also represent the sequence of presentation of the practice block, the two baselines blocks and the two FAST blocks.

FAST Block Procedure

The presentation of stimuli and feedback, criterion, and maximum trial limit during all blocks was almost identical to the *General FAST Block Procedure* described in Experiment 1. However, the response window was shortened to 2000ms for reasons explained in the Introduction. Further, the number of current number of correct successive responses was continuously displayed at the bottom of the screen, again as explained in the Introduction. Lastly, a practice block was also introduced, as explained below.

Practice Block

The newly devised practice block was the first block to be administered in the new FAST procedure. This block utilized novel nonsense syllables (P1, P2, Q1, and Q2) and was identical in format to a baseline block (described below). Participants were presented with the following instructions page at the onset of the Practice Block.

THE FOLLOWING SECTION IS A **PRACTICE BLOCK**

IN THE FOLLOWING SECTION

A single word will appear on screen

For Example

“bix”

Each time you see the word you must press either

The “**Z**” key or the “**M**” key

After you press a key, the computer will tell you if you were

CORRECT or WRONG

If you don’t press anything within 2 seconds of seeing a word, the computer will count this as a WRONG answer and move on to the next word

Your task is to learn which key **(Z or M)** to press for each of the words. This section will continue until you can make 10 CORRECT RESPONSES in a row WITHOUT ANY MISTAKES

Instructions remained fixed on screen for a period of 10s, after which time the following appeared at the bottom of the screen in 30pt font; "When you are ready to start the next section, Press Any Key". Instructions remained on screen until the participant pressed a key, which started the practice block.

After the block was completed by the participant, a screen appeared with the following text:

Well done! You have completed this section of the experiment.

Please take a short break to rest your eyes from the computer screen.

The instructions for the next section will be displayed shortly.

This text remained on screen for 30s, after which the instructions for the next block were automatically displayed

Baseline Blocks

Two baseline blocks were presented (one before the first FAST testing, and one after the final FAST block). As usual, Baseline blocks involved novel and unique stimuli. Baseline 1 employed X1, X2, Y1, Y2 as stimuli, while Baseline 2 employed X3, X4, Y3, Y4 as stimuli. These phases attempted to establish the following relations, respectively; X1-Y1, X2-Y2 and X3-Y3, X4-Y4.

At the onset of the block, the following instructions were displayed:

IN THE FOLLOWING SECTION

A single word will appear on screen

For Example

“bix”

Each time you see the word you must press either

The “**Z**” key or the “**M**” key

After you press a key, the computer will tell you if you were

CORRECT or WRONG

If you don’t press anything within 2 seconds of seeing a word, the computer will count this as a WRONG answer and move on to the next word

Your task is to learn which key **(Z or M)** to press for each of the words. This section will continue until you can make 10 CORRECT RESPONSES in a row WITHOUT ANY MISTAKES

Instructions remained fixed on screen for a period of 10s, after which time words “When you are ready to start the next section, Press Any Key” appeared at the bottom of the screen in 30pt font. When the participant pressed a key, the block began.

After the first baseline block was completed by the participant, a screen appeared with the following text:

Well done! You have completed this section of the experiment.

Please take a short break to rest your eyes from the computer screen.

The instructions for the next section will be displayed shortly.

This text remained on screen for 30s, after which the instructions for the next block were displayed. When the participant completed the second baseline block (at the end of the experiment) text appeared thanking the participant for their participation and instructing them to contact the researcher.

Test Blocks

The test blocks utilized the B1 and C1 stimuli used in the MTS training, and two further novel nonsense syllables, N1 and N2. The first of these blocks (consistent) established two functional response classes (C1-B1 and N1-N2) that were consistent with the derived relations predicted given the MTS training. The second test block established two functional response classes that were inconsistent with the relations trained in Phase 1 (B1-N1 and C1-N2). Unlike in previous experiments, the block order was fixed as consistent first (see Introduction.)

The instruction page displayed at the start of each block was identical to the one presented for the baseline block. Once again, instructions remained fixed on screen for a period of 10s, after which time words "When you are ready to start the next section, Press Any Key" appeared at the bottom of the screen in 30pt font. Instructions remained on screen until the participant pressed a key, which in turn began the block.

After the block was completed by the participant, a screen appeared with the following text:

Well done! You have completed this section of the experiment.

Please take a short break to rest your eyes from the computer screen.

The instructions for the next section will be displayed shortly.

This text remained on screen for 30s, after which the instructions for the next block were displayed

Results

Matching to Sample Training and Testing

Five participants completed all stages of equivalence training and the equivalence test. The two participants who did not complete all stages were unable to commit to the final day of the procedure within an appropriate time frame, and so were dropped from the study. Most participants required only one block to complete each stage of training. S2 required two blocks to complete stage 8 (in which participants required 31/32 correct responses to proceed, with 50% reinforcement and a six second delay between observation response and presentation of comparisons.) S4 required two blocks to complete stage 2 (in which 31/32 responses were required, with 100% reinforcement and no delay.) S7 required six blocks to complete stage 1 training and five to complete stage 2, suggesting some difficulty in initial learning of the required relations. However, no participants required more than one block to complete the equivalence testing block successfully. Table 1 shows the number of blocks required for each participant to complete each stage of the MTS training and testing.

Table 1: Number of blocks to criterion for each participant at each stage of the training phases. (Note that different blocks have different lengths and criteria, see method)

Sub	Pre-train	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9	Test
1	1	1	1	1	1	1	1	1	2	1	1
2	1	1	2	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1
5	1	3	1	1	1	1	1	1	1	1	1
7	1	6	5	1	1	1	1	1	1	1	1

Function Acquisition Speed Test

Each of the five participants who completed the MTS training and testing proceeded to complete the FAST procedure. Table 2 shows the trials to criterion for each participant on each block of the FAST, as well as Mean Baseline scores and Strength of Relation indices

Table 2: Trials to Criterion in Practice Blocks, Baseline Blocks (Base1 and Base2), Consistent Blocks (Con) and Inconsistent Blocks (Incon). Also shown are Mean Baselines, and SoR indices for each participant.

Sub	Practice	Base 1	Base 2	Mean Base	Con	Incon	SoR
1	88	14	13	13.5	16	30	5.379
2	43	19	29	24	24	32	2.517
4	20	11	13	12	11	21	4.024
5	85	53	36	44.5	30	63	8.694
7	123	52	134	93	99	73	-5.736
Mean (SD)	71.8 (40.53)	29.8 (20.92)	45 (50.75)	37.4 (33.68)	36 (35.97)	43.8 (22.7)	2.975 (5.38)

Four out of five participants showed a FAST effect in the predicted direction. The mean positive SoR score was 5.154 (SD= 2.634), noticeably higher than seen in previous experiments. This indicates that the more extensive and robust MTS training resulted in stronger relations, and subsequently stronger FAST effects for the majority of participants.

Discussion

Over the course of three days, in sessions lasting approximately 30 minutes, participants completed nine stages of progressively more challenging matching to sample training. This training procedure was designed to maximise the probability that stable equivalence relations would emerge. After completing this training, the participants completed a FAST procedure. Four out of five participants showed a strong positive SoR index, demonstrating that when equivalence training is robust, the FAST is 80% effective at detecting the derived relations learned by the participants in pre-experimental training. This result is very important in the overall context of the thesis, as it supports the suggestion that a significant amount of negative results (i.e. SoR of less than zero) could be attributed to insufficiently trained equivalence relations in the pre-FAST training phases. Even when participants pass an equivalence test, the change in format and contingencies may be enough to disrupt those newly formed equivalence relations sufficiently to produce a negative SoR index.

The nature of the experimental design for the current study necessitated a small sample size due to the arduous nature of the training and testing protocol. Participants were required to attend three sessions of MTS training, with the FAST test on the third session. These sessions occurred over a period of 3-5 days, with a minimum of 24 hours between the beginning of one session and next, and a maximum of 48 hours. While this sample is small, it should be remembered that small n research with extensive experimental control is a mainstay of behaviour analytic research (Sidman, 1974). The reader should view the five performances outlined here as replications of each other and the main effects, rather than as random samples in a statistically oriented survey (e.g., Yin, 2003).

The single participant who did not show a FAST effect in the expected direction (S7) showed the highest numbers of trials to criterion on every block of the FAST expect for the initial Baseline phase (in which their score was only 1 lower than the highest recorded). This suggests that participants who have difficulty with the task *in general* may not reliably show FAST effects. In the pre-training phases, S7 also showed the highest result on any single block, requiring 6 cycles on the first training block. This may indicate difficulty in adapting to a new procedure and new response requirements. Also of note was that this participant was significantly older than other participants (26 years older than the second oldest participant). Age and cognitive ability have been known to have an effect on implicit test outcomes (Greenwald et al., 2003) and it may be that the FAST suffers from similar issues.

With regard to the modifications to the FAST procedure, these appear to have had a positive effect on FAST outcomes. The reduction in the response window did not lead to any participants having unusual difficulty in completing the procedure successfully. While S7 did display some difficulty, it was not out of line with participants who have similarly struggled with a 3s response window. The question of response windows will be further addressed in Chapter 3 of this thesis, when procedural modifications will be studied more systematically.

As regards the fixed block order, while this modification was appropriate in a circumstance in which the experimenter is aware of the relations under examination, when the relations under investigation are real world relations, it is not possible to *a priori* assume which relations are in fact “consistent”. As such, a fixed block order was not retained for the experiments in Chapter 4.

The practice block and the counter showing the participants’ current progress towards criterion are both positive developments. Informal debriefing with participants suggest that participants found both of these innovations useful.

Experiment 5: Using the FAST to measure untested derived relations

After the strong results shown in Experiment 4, the basic functionality of the FAST is well established. Challenges still remain with regard to fine tuning the experimental procedure to maximise the stability and validity of the test outcomes, but in light of the experiments in this chapter, it can be reasonably concluded that the basic FAST methodology is an avenue of research worth exploring in detail in this thesis and beyond.

An exciting challenge for the FAST procedure is to determine whether or not it can detect relations that have never even been derived by the participant following training that merely “implies” those derived relations. In experimental terms, what is needed is the demonstration of a FAST effect before equivalence testing has even taken place. This is a crucial challenge for any implicit test as it is the only way to prove that the test is capable of detecting relations implicit in a participant’s history but which have never been consciously derived or tacted. If this can be achieved, it would suggest that social histories may be capable of interfering with further learning by virtue of implied stimulus relations of which individuals themselves are not even aware. The author uses the word “untacted” to refer to such relations, as the participant is not in psychological contact with the history of derived relations that is controlling their behaviour, as such is unable to report the source of this behaviour. As such, it is an extension of the concept of the “tact” – a verbal operant in which a person names a stimulus with which they are in contact through the senses.

For example, a history in which the word “Black” is related to the word “African’, which is in turn related to the word “lazy”, may impede learning an association between a black person’s face and positive words describing the nice sweater they are wearing. In other words, even following attempts to teach an individual something positive regarding a particular Black person, the acquisition of the required stimulus relations (e.g., *Black person-nice*) may be slower than if the relevant stimuli did not participate in any conflicting implied (untacted derived) relations (e.g., *Black person – not nice*). Thus, an individual’s history may create resistance to such learning, perhaps without the individual ever being able to report why that is the case. Other psychologists may refer to this as an “unconscious process”, but given our current functional-analytic approach, we are able to keep the process of learning interference overt and open to direct experimental analysis.

Method

Participants

Thirty-eight participants (12 males and 26 females) participated in the experiment. The participants’ ages ranged from 18 to 54 (Mean = 21.82, SD = 5.54). Twenty-eight participants were exposed to the sequence of Equivalence Training, FAST, Equivalence Testing. 10 participants were exposed to the sequence of Equivalence Training, Equivalence Testing, FAST (see below). Of the 28 participants in group 1, 11 participants failed to complete the equivalence test phase successfully and were classified as control participants (see Results).

Apparatus and Stimuli

All phases of the experiment were presented to participants on an Apple Macbook laptop using the Psyscope software package, as per Experiment 1. Stimuli consisted of twenty-eight nonsense syllables (see Appendix 2) arbitrarily assigned to their roles as samples, comparisons, and FAST test stimuli (see below). These will be referred to hereafter using alphanumerics.

General Experimental Procedure

The experiment consisted of three phases. Phase one (Equivalence Training) consisted of a matching-to-sample protocol to establish two 3-member equivalence classes, each containing three nonsense syllables (A1-B1-C1 and A2-B2-C2). After completion of Equivalence Training, participants were randomly assigned to one of two experimental groups. Group one were exposed to the Equivalence Testing Phase followed by the FAST

Phase. Group two were exposed first to the FAST Phase and then the Equivalence Testing Phase. The Equivalence Testing Phase tested for the presence of the derived relations that emerged from Phase 1 training (i.e., A1-C1, C1-A1, A2-C2, and C2-A2). The FAST phase consisted of three runs (exposures) of the FAST test block pairs, with baseline blocks administered between them. Multiple runs of the FAST were employed so that the stability of participant performance across time could be considered.

Phase 1: Equivalence Training

In this phase, participants were exposed to a linear matching-to-sample procedure designed to establish two 3-member equivalence classes according to a one-to-many (sample-as-node) protocol. The relations trained in this phase were identical to those trained in the Equivalence Training phase of Experiment 3 (see also Figure 10, below).

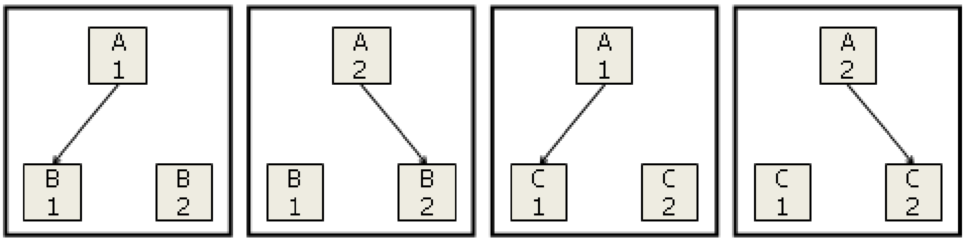


Figure 10. The stimulus relations trained during Phase 1. The arrows indicate reinforced matching responses.

Participants were presented with the following instructions at the onset of Phase 1:

In a moment some words will appear on this screen. Your task is to look at the word at the top of the screen and choose one of the two words at the bottom of the screen by "clicking on it" using the computer mouse and cursor. During this stage the computer will provide you with feedback on your performance. You should try to get as many answers correct as possible. If you have any questions please ask them now. When you are ready please click the mouse button.

Trial presentation was the same as in Experiment 1, with the exceptions noted above. Participants were required to complete a block with 30/32 correct responses. If participants failed to meet criterion, the training block was repeated until criterion was reached.

Phase 2: Equivalence Testing

During equivalence testing, probes for the unreinforced formation of B-C relations (transitivity) and C-A relations (combined symmetry and transitivity) were presented (see Figure 11, below). The testing format was similar to the training procedure. However, no corrective feedback was provided. Again, participants were required to reach a criterion of 30/32 correct responses. If participants failed to reach criterion after 4 testing blocks, they were classified as control participants and they proceeded to Phase 3 as normal.

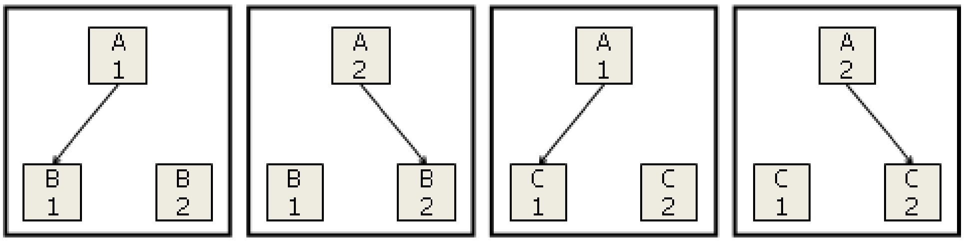


Figure 11. The derived stimulus relations probed for during The Equivalence Testing Phase. Arrows indicated correct choices.

Function Acquisition Speed Test

The FAST consisted of a practice block, two test blocks and two baseline blocks. One baseline block was administered before the test blocks, and one afterwards. The FAST procedure used in this experiment was identical to that used in Experiment 4. A schematic of the block order can be seen in Figure 12 below.

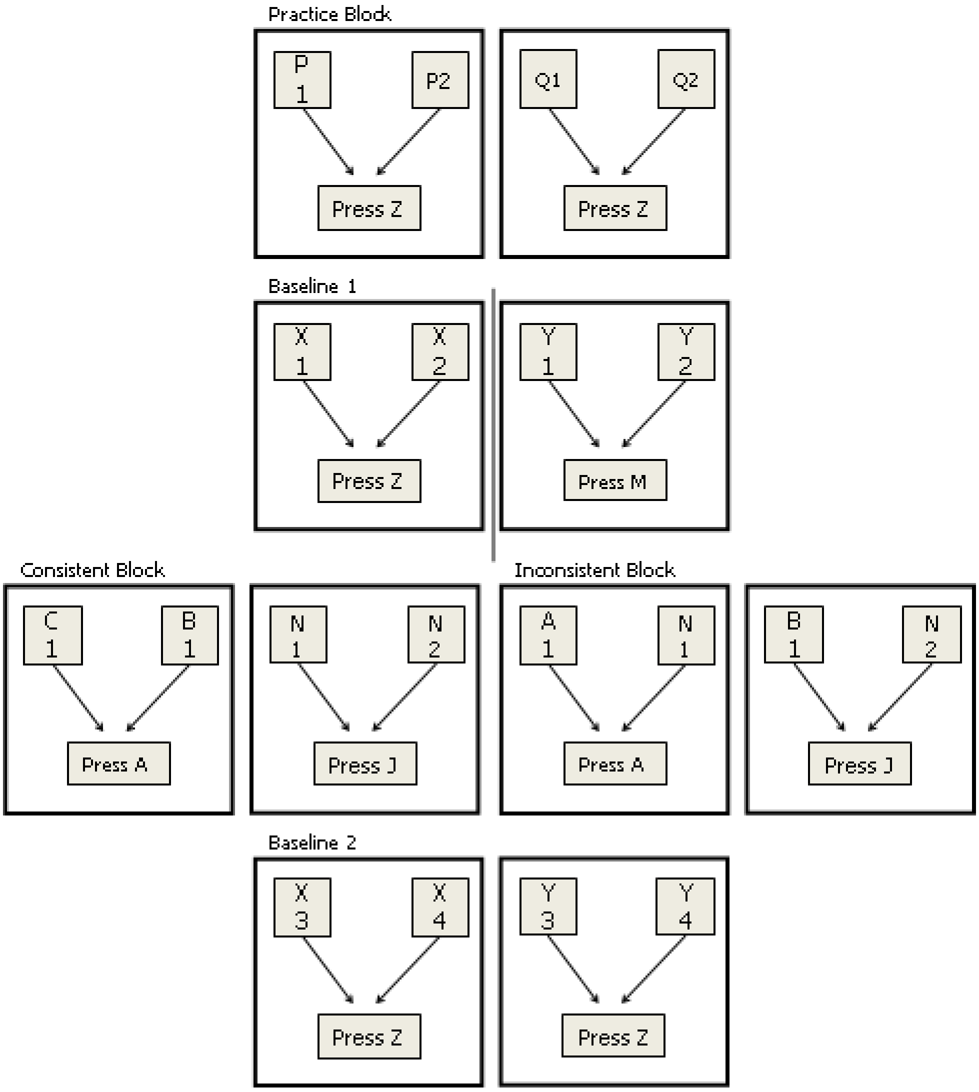


Figure 12. *A schematic of the various response functions assigned to stimuli during practice, baseline and FAST blocks. The four rows, read from top to bottom, also represent the sequence of presentation of the practice block, the two baselines blocks and the two FAST blocks.*

Results

Matching to Sample Training

All participants completed matching to sample training. The average number of training blocks required to reach criterion (97% correct) was 2.79. (SD=1.803). Table 1 illustrates the training requirements for each participant.

Table 1: The number of equivalence training blocks required for each participant to reach criteria. Also note that the first two columns comprise Group 1 participants, the third Group 2, and the fourth Group 3.

Participant	Blocks Required	Participant	Blocks Required	Participant	Blocks Required	Participant	Blocks Required
1	3	10	2	18	4	29	3
2	2	11	2	19	8	30	2
3	2	12	2	20	4	31	4
4	4	13	4	21	2	32	4
5	1	14	4	22	2	33	1
6	2	15	1	23	2	34	2
7	2	16	1	24	3	35	2
8	10	17	2	25	2	36	2
9	2			26	5	37	2
				27	2	38	2
				28	2		
Mean			5.111		3.272		2.4

Equivalence Testing

All participants completed an Equivalence Testing phase. Participants in group 1 received the equivalence testing phase after the completion of the FAST block. 10

participants failed to complete the testing phase successfully. These participants were separated out for the purposes of the rest of the data analysis.

FAST Practice Blocks

All participants completed a practice block prior to the FAST. The trials to criterion for each participant are displayed below on table 2.

Table 2: The trials required to reach criterion on the practice block for each participant. Also note that the first two columns comprise Group 1 participants, the third Group 2, and the fourth Group 3.

Participant	Trials	Participant	Trials	Participant	Trials	Participant	Trials
1	72	10	23	18	26	29	43
2	200	11	63	19	39	30	23
3	41	12	17	20	109	31	26
4	174	13	20	21	17	32	84
5	16	14	200	22	93	33	137
6	70	15	113	23	88	34	10
7	39	16	60	24	40	35	29
8	46	17	65	25	33	36	19
9	37			26	57	37	26
				27	21	38	35
				28	79		
Means			73.88		54.727		43.2

FAST Baseline Blocks

Table 3 shows the number of trials required for each participant to reach the fluency criterion (10 correct responses in a row) for each of the two baseline blocks, as well as the mean baseline performance for each participant. A related samples Wilcoxon Signed Ranks test showed that the differences from Baseline 1 to Baseline 2 was statistically significant (z= -3.187, p=0.001), suggested a practice effect across Baselines.

Table 3: Trials to criterion for each participant for the two Baseline Blocks, and the Mean Trials to Criterion for each participant.

Participant	Base 1	Base2	Mean	Participant	Base 1	Base 2	Mean
1	13	45	29	10	12	12	12
2	115	30	72.5	11	17	16	16.5
3	23	32	27.5	12	68	19	43.5
4	21	25	23	13	26	26	26
5	17	12	14.5	14	75	10	42.5
6	23	14	18.5	15	16	117	66.5
7	13	13	13	16	36	17	26.5
8	35	12	23.5	17	139	68	103.5
9	31	13	22				
Participant	Base 1	Base2	Mean	Participant	Base 1	Base 2	Mean
18	44	18	31	29	15	12	13.5
19	44	12	28	30	52	38	45
20	36	45	40.5	31	84	10	47
21	18	10	14	32	75	73	74
22	170	14	92	33	33	24	28.5
23	48	17	32.5	34	83	19	51
24	12	19	15.5	35	47	18	32.5
25	89	32	60.5	36	43	31	37
26	10	25	17.5	37	32	30	31
27	13	20	16.5	38	46	34	40
28	30	16	23				

Test Blocks

Each participant completed two FAST test blocks, a consistent block and an inconsistent block. The primary datum for the FAST is the difference in trials to criterion between consistent and inconsistent blocks. The sections below describe the number of trials to criterion for each participant and the order in which the blocks where administered. (See Tables 3, 4 and 5, below) A Mann Whitney U test showed that there was no order effect across all participants (z=1.454, p=0.152).

Group 1. The first experimental group consisted of the 17 participants who completed the FAST before completing equivalence testing, but also passed the later equivalence test. Of those 17 participants, 9 showed a difference in the expected direction. Table 4 shows the number of trials to criterion for each participant on each test block, and the difference between consistent and inconsistent blocks. The mean differential between blocks was -.41 (SD=12.699) for this group. An SoR index was also calculated for each participant (see Table 4). The mean SoR index was -0.32 (SD = 3.78). A one-sample t test indicated that the SoRs were not significantly different to zero (t=- 0.348, p = .732). These results indicate that when presented before an equivalence test, the FAST procedure is unreliable in its ability to detect the history of matching to sample training.

Group 2. The second group consisted of the 10 participants who completed the FAST before completing equivalence testing, but who failed the later equivalence test. Of those 10 participants, six displayed a negative SoR index, and four displayed a positive SoR index. Table 5 shows the number of trials to criterion for each participant on each test block, and the difference between consistent and inconsistent blocks. The mean differential between blocks was -.1.82 (SD=25.031) for this group. An SoR index was also calculated for each participant (see Table 5). The mean SoR index was -0.6822 (SD = 6.968). A one sample t test shows that these SoR indices were not significantly different from zero (t=-0.325, p= .752). This result is to be expected given that participants did not successfully pass equivalence testing. On a group level, their performances on the FAST were highly variable (with a higher SD than the "pass" group above).

Group 3. The third experimental group consisted of the 10 participants who completed the FAST after completing equivalence testing, but also passed the later equivalence test. Of those 10, 7 showed a difference in the expected direction (i.e. these participants took longer to reach criterion on the inconsistent block compared to the consistent block). Table 6 shows the number of trials to criterion for each participant on each test block, and the difference between consistent and inconsistent blocks. The mean differential between blocks was 11.4 (SD=21.864) for this group. An SoR index was also calculated for each participant (see Table 6). The mean SoR index was 2.913 (SD = 5.135). A single sample t-test indicated that the SoR scores were significantly different from zero (t=1.794, p=.05). These results suggest that when a participant has been required to tact derived relations (albeit without reinforcement) the FAST is sensitive to that history.

Table 4: SoR Indices for each participant. Also note that the first two columns comprise Group 1 participants, the third Group 2, and the fourth Group 3

Participant	SoR	Participant	SoR	Participant	SoR	Participant	SoR
1	-7.42	10	.40	18	3.20	29	0
2	5.37	11	-7.13	19	-2.10	30	2.36
3	-3.92	12	-.80	20	-1.08	31	-.78
4	-4.47	13	3.38	21	3.03	32	16.03
5	2.62	14	2.40	22	.44	33	2.09
6	-1.03	15	1.43	23	-18.1	34	-3.05
7	-.39	16	4.27	24	.36	35	4.88
8	-3.17	17	1.08	25	10.48	36	1.38
9	1.94			26	2.45	37	3.49
				27	-4.28	38	2.71
				28	-1.91		

Mean			-.32		-.68		2.911

Discussion

In the current experiment, participants who received a FAST before the presentation of an equivalence test, as well as those who failed an equivalence test presented afterward a FAST did not show any significant FAST effect on the group level. However, when participants were presented with the FAST after an equivalence test (which they passed), the majority of participants showed a FAST effect, and the effect was significant at a group level.

These results demonstrate that the FAST is sensitive to derived relations, but not necessarily those that have not yet been explicitly derived and tacted by the participant. An equivalence test establishes contingencies under which participants must tact the derived relations implied by the MTS training procedure. The current results suggest that until the participant has had the opportunity to tact derived relations in a matching-to-sample context, these relations do not produce a detectable perturbation in the learning of new relations that could be detected by the FAST. This is a fertile avenue for future research to explore.

General Discussion

The relational frame theory account of attitudes describes an attitude as *a network of derived and explicitly reinforced stimulus relations according to which the functions of events are transformed.* As such, the existence of a given attitude in a person is the result of a particular history of verbal interaction with a socio-verbal community, and that history is brought to bear on new events and transforms them accordingly (Barnes & Roche 1997, Dixon et al., 2003; Grey & Barnes, 1996,).

Given this conception of attitudes, any measure that aims to detect attitudes must be sensitive to the effects of that history. The Watt et al. (1991) paradigm and related research (Leslie et al., 1993, Merwin & Wilson, 1995; Moxon et al., 1996; Plaud et al., 1995) demonstrated that previously established stimulus relations interfered with the acquisition of novel stimulus relations involving stimuli with histories in competition with the relations being trained. The RFT model described above also indicates that the constituent elements of an attitude are stimulus relations. These two ideas, taken together, informed the conception of the Function Acquisition Speed Test methodology and the early research strategy. Research was directed towards building the methodology such that it was capable of detecting these elemental stimulus relations under controlled laboratory conditions.

Each experiment in Chapter 2 used the same research strategy. Participants were trained to relate a number of stimuli in the laboratory. Immediately thereafter, participants completed a FAST procedure in which the test stimuli where two of the stimuli that had been related in the pre-FAST training.

In Experiment 1 twnty one participants were trained using a matching to sample procedure to produce to A-B relations (A1-B1 and A2-B2). The stimuli used were three letter nonsense syllables, with which the participants had no previous experience. After participants had completed the matching to sample procedure, they were required to complete a FAST.

The FAST consisted of a Baseline Block, two Test Blocks (Consistent and Inconsistent) and a second Baseline Block. Each block of the FAST uses four stimuli, and participants must learn to respond with either a left key press (e.g. press z) or a right key press (e.g. press m) within a 3s response window when presented with a stimulus. Two stimuli require a right key press, and the other two require a left key press. The instructions do not indicate which is the correct response to each stimulus – participants must learn via corrective feedback alone. In order to complete a FAST block, participants were required to make 10 correct responses in a row without error.

The Baseline Blocks utilized four novel nonsense syllable stimuli each. FAST Test B1ocks presented two novel nonsense syllable stimuli, and two *test stimuli* – the stimuli whose relatedness is under investigation. In Experiment 1, the test stimuli were the A1 and B1 stimuli from the pre-FAST training phase. In the consistent block A1 and B1 both required the same response, while in the inconsistent block A1 and B1 required different responses. The number of trials-to-criterion required by each participant is the basic datum of the FAST, and is used to calculate a Strength of Relation index (SoR = Inconsistent –Consistent /Mean Baseline). The FAST methodology predicts that it will be more difficult to complete the inconsistent block if the two test stimuli have a history of being related, which yield a positive SoR index – a positive FAST effect. Thirteen out of the eighteen participants who completed the procedure took a greater number of trials to complete the inconsistent block compared to the consistent block, showing a positive SoR index. Further the five negative SoR results were noticeably smaller than the positive SoRs. However, there was some concern with regard to the high level of within participant variability on Baseline Blocks, though there was no statistically significant practice effect.

Experiment 3 sought to extend the success of Experiment 1. In the stimulus relation training phase, twenty-two participants completed a linear matching to sample protocol to establish two three member equivalence classes (A1-B1-C1 and A2-B2-C2). They were then tested for the emergence of the derived relations A1-C1 and A2-C2, before proceeding to the FAST procedure. Participants completed three consecutive FAST procedures in order to assess the stability of the effect over multiple exposures. Further, the criterion for completion of a block was changed from ten consecutive correct responses to nine correct responses in any sequence of ten, allowing one error or non-response.

The FAST once again used novel nonsense syllable stimuli in the Baseline Blocks and as two of the stimuli in the Test Blocks. The Test Stimuli were A1 and C1 from the previous phase. In each of the runs of the FAST, a majority of participants showed a FAST effect in the expected direction. The method of calculating the SoR was altered slightly (The mean baseline score was transformed via natural logarithm, in order to create a narrower range of larger SoR scores). On the first and third runs of the FAST, group level FAST effects were significantly different from zero, while on the second run, the difference approached significance ($p = .08$). The results confirmed that the FAST procedure was able to detect derived relations in a majority of the test takers. However, concerns were raised with regard to the stability of the results and the number of false negatives. It was suggested that some of this unreliability was due to the nature of the equivalence training procedure. The equivalence relations being tested were only recently established, and as such, the disruptive contingencies in the inconsistent phase may have been sufficient to override the pre-FAST training.

Experiment 4 was designed to address some of those issues. The pre-FAST training phase of the experiment, which again used nonsense syllables as stimuli, was designed to maximise the probability of equivalence relations emerging in the five participant, and to ensure that these relations were as robust and well established as possible. The equivalence training procedure changed to a One-To-Many protocol (Arntzen et al., 2010), and participants were required to emit an observation response before comparison stimuli were displayed (Arntzen et al., 2011). Participants completed nine increasingly challenging stages of training over 3 days. As the stages progressed, the delay between observation response and display of the comparison stimuli was increased incrementally (Arntzen, 2006), and a fading of consequences was introduced (Arntzen et al., 2010). The criteria necessary to complete each phase of the training increased as the stages progressed. Due to the extensive time required from each participant, a repeated single participant design was used.

Further, the FAST procedure was modified; the response window was shortened to 2000ms, a practice block was included, the instructions were changed to increase clarity, and a “counter” was included on the screen during the test phases which displayed the number of correct responses in a row that participants had achieved. Four out of five participants showed strong positive SoR indices on the FAST.

The results of Experiments 1, 3, and 4 combine to establish proof-of-concept for the FAST procedure. Being able to detect both basic stimulus relations and derived stimulus relations is vital for a behavioural implicit test, as these are considered to be the constituent relations of an “attitude” (Grey & Barnes, 1996). It must be borne in mind that the positive results in Experiments 1 and 3 occurred in the context of only recently established stimulus relations. Even with the rigorous training used in Experiment 4, these relations cannot hope to approach the level of rehearsal and reinforcement that the real world relations underlying attitudes have in a socio-verbal community. As such, that the FAST detects these relations across multiple experiments is highly promising, as the real world relations for which the experimentally trained relations are an analogue will be significantly more entrenched in the participants’ behavioural history.

However, not all of the initial experiments can be considered a complete success. In Experiment 2, eleven participants completed a function training and testing protocol which established an erotic function for a nonsense syllable A1 and an aversive function for nonsense syllable A2 by pairing these stimuli with images. Participants then completed three consecutive FASTs, in which the test stimuli were A1 and three novel erotic images which were displayed quasi randomly. The other two stimuli used where novel nonsense syllables.

The results proved to be surprising, as only two participants showed the expected result on the first run of the FAST. Eight participants showed a positive SoR on the second run, with their SoRs being statistically different from zero on the group level. The IAT model of implicit attitudes strongly suggests that implicit attitudes have a strong evaluative component. This is not supported by the evidence from this experiment, and evidence from other experiments (particularly Experiment 5) suggests that only consciously tacted relations are detectable by a FAST procedure. However, this cannot be concluded concretely, due to issues with experimental design.

The FAST effect faded from statistical significance again on the third run, with only six participants showing a positive SoR. It was suggested that this may be due to insufficient training in the function training phase. However, it is more likely that these unstable results were due to the use of multiple image stimuli in place of the second test stimulus. The use of multiple exemplars, and a novel stimulus type, may have been disruptive to participants in and of itself. In brief, Experiment 2 overreached from the basic FAST procedure, and this modification was found to be ineffective and unreliable. The use of image stimuli and multiple exemplars will be returned to in the Chapter 4, and future FAST research should return to the question of the role played by evaluative conditioning in attitudes with the benefit of later improvements that were made to the FAST procedure.

Experiment 5 represented an extension of Experiments 3. This experiment used the same training procedure as Experiment 3, but with one important difference – One group of ten participants completed a FAST *before* completing Equivalence Testing, while another twenty-eight completed Equivalence Testing prior to the FAST (as in Experiment 3). The latter group showed the expected positive FAST effect. However, neither the group who passed the later equivalence test nor those who failed it showed a positive SOR. This experiment is valuable in that it demonstrates the limits of the FAST procedure, and suggests that perhaps untacted derived relations do not have an effect on subsequent learning of relations in the same way that tacted relations do. Of course, once again, the brief nature of the pre-training may have obfuscated subtle differences between the pass/fail groups who received a FAST first. A natural extension of this experiment would involve extensive training to form equivalence classes (as in experiment 4), with the equivalence testing likewise positioned after the FAST procedure. This is an exciting research direction for the FAST. Aside from its potential use as an implicit test in applied contexts, the FAST methodology may be useful research tool for exploring the RFT model of attitudes and refine its ability to explain the “implicit cognition” phenomenon. Discovering the conditions under which a history which has a high probability of emergent derived relations, but that has not been tested (i.e. the relations are untacted, or “implicit”), is sufficient to affect future behaviour on a FAST would add an additional layer of detail to the RFT model, and expand its explanatory power to include attitudes which are outside the conscious awareness of the participant.

It is important to confront the matter of large variances in FAST effects across participants and across runs, particularly in experiments 2 and 3, where multiple runs of the FAST were utilised. While group trends were generally in the expected direction, even on a third repeated exposure to the FAST, large variances in performances within and across participants are indicative of less than perfect experimental control.

The practice effect observed for all participants across the several baseline phases offers a clue as to one likely source of such variance. More specifically, the observation of variance in baseline performances within participants across time allows us to separate sources of control related to the experimental stimuli (i.e. during FAST blocks) from those related only to the FAST procedure itself. That is, this observation confirms that variability in performances are observed even when the stimuli involved are novel and randomly selected. Thus, it would appear that the acquisition rates being measured here by the critical FAST blocks are themselves variable and much or maybe most of this variability is controlled by sources separable from the derived relational responding contingencies.

The different levels of pre-FAST training across these experiments contribute to our understanding of the behavioural variability observed in FAST effects across participants. That is, during the inconsistent test block, responses which ran counter to the equivalence training were reinforced. Across multiple runs, or in instances where the participant was exposed to the inconsistent block first, this may have been sufficient to destabilize equivalence relations, especially if these were poorly trained. In effect, the very relations on which the FAST effect depends may have been destabilized during the course of the FAST itself, leading to varying and poorly understood outcomes that changed rather than stabilised across exposures. This issue will be further examined in the following chapter, in which multiple runs of the FAST are utilized in all experiments.

Chapter 3

Exploring procedural modifications of the FAST methodology

In the previous chapter, the FAST methodology has demonstrated its efficacy in detecting experimentally trained stimulus relations, and in exploring naturally occurring stimulus relations. As that research progressed, small procedural modifications were made for various specific reasons and sometimes purely for the purposes of conducting good inductive work. For instance, in Experiment 3, the response criterion was changed from 10 correct responses in a row to 9/10 correct responses in any consecutive 10 trials. While these changes were introduced in the pursuit of a basic experimental effect, a more systematic experimental investigation of the effects of such procedural changes is now merited. It may be that even small changes in the contingencies (in the case of response windows and block completion criteria) or in the participants' history with similar tasks (in the case of practice blocks) might have profound effects on the behaviour produced, and as such, we must understand the effect of different experimental preparations on the behaviour of interest.

This chapter represents a more focused analysis of test feature manipulation. Each of the following experiments explores the effect of varying different aspects of the FAST procedure on the resulting "FAST effect". The reader will have noted that a number of changes were introduced to the FAST procedure between Experiments 4 and 5. The experiments described in the current chapter were performed before Experiment 4 and 5, and the changes to the procedure described therein were partially as a result of the findings in this chapter.

Experiment 6 explores the effect of different response windows on FAST performance. Experiment 7 explores the effect of different numbers of practice blocks. Experiment 8 investigates the effect of varying the fluency criterion for completing a FAST block. In all experiments in this chapter, participants were trained in simple A-B relations (as in Experiment 1) using nonsense syllables, and then subsequently were asked to complete a series of FAST procedures to probe for the trained relations. Each experiment uses a similar between-groups design and controls fully for all test features bar the one of interest. Multiple runs of the FAST are utilized to explore various effects over multiple exposures. These experiments are intended to guide the researcher towards the optimal FAST procedure and maximise its sensitivity and stability.

Experiment 6: Response Windows

The majority of implicit testing methodologies use response time differentials as their primary metric for inferring the existence of implicit attitudes. The IAT scoring algorithm (Greenwald et al., 2003) recommends that response latencies on individual trials which exceed 3000ms are truncated to 3000ms for the purposes of data analysis. The IRAP used the same method in its earlier experiments. (e.g., Barnes-Holmes et al., 2006; Power, Barnes-Holmes, & Barnes-Holmes, 2010). However, the IRAP later altered its approach. Participants are currently required to complete a minimum of two practice blocks. The participants are required to complete the two blocks with a median response time of 2000ms in order to proceed to the test blocks. Initially, the criterion was 3000ms, but research concluded that reducing the criterion to 2000ms increased reliability and validity (Barnes-Holmes, et al. 2010). Further, the current IRAP procedure includes the feedback "Too Slow" when participants fail to respond within 2000ms, although they are still afforded the opportunity to respond. When participants fail to maintain a mean latency of 2000ms in any of the test blocks, the data for that participant is discarded (Barnes-Holmes, et al. 2010). Both the IRAP and IAT procedures require response latencies to fall within a tight window (either in vivo during task performance, or post-hoc vis-à-vis data manipulation) in order to indicate implicit cognitive associations. Both methods aggressively remove outliers to maintain a data set which falls within certain response time parameters. It is impossible for critics to know how much data is being lost via these types of procedures without scrutinizing raw data. Moreover, most IAT software actually records latencies of over 3000ms as 3000ms and so records of actual response time distributions are never even recorded in the first instance (this practice is testimony to the relative lack of interest social cognitive researchers have in the actual behaviour of their participants).

The FAST takes a different approach, by allowing participants a response window in which to respond, and counting failures to respond within the window as an incorrect response. In this way no responses are discarded as non-responses (a misnomer within behavioural philosophy) or recoded. This approach is similar to the one taken in the behaviourally modified IAT (Gavin et al., 2008). The IRAP uses a pseudo-response window, in that it requires participants to complete practice blocks until their average response time falls below a 2000ms criterion (Barnes-Holmes et al., 2010). The FAST uses trials-to-criterion as its primary metric, and its criterion is fluency based. In order for a participant to be considered fluent with a response, responding must be both consistent and rapid. As such, changing the response window of the FAST represents a change in the criterion required to complete the task. A faster response window would necessitate more fluent responses from participants and may lead to different performances.

Time pressure has also been implicated in producing the effects that underpin the IAT and IRAP, and all other response time based measures (REC model: Barnes-Holmes et al., 2010). As such, it is in the interest of implicit tests to require the fastest possible responding from their participants, so as to mitigate effects from more complex verbal behaviour, such as rule following, on the test outcomes. However, an overly stringent response window could render it too difficult for participants to produce a series of correct responses at all, thus increasing the number of participants for whom “time outs” are recorded. Moreover, a large number of missed responses may or may not make data difficult to interpret. In any case, it is important to try to identify the boundary conditions for the FAST effect in terms of response-window durations of varying length. The current experiment seeks to explore the effects on FAST performance of varying the response window across four different levels.

Method

Participants

Forty four volunteers participated in the study as participants. Data from participant 32 was excluded due to technical errors in the presentation of the FAST. Twelve participants failed to complete FAST blocks within the 100 trial limit on at least one occasion. Of the remaining 31, 12 were male and 20 were female, aged 19-25 (Mean 21.23, SD =1.04). The eliminated participants have been excluded from the initial data analysis. However, their data will be considered further in the Results section.

Apparatus and Stimuli

All phases of the experiment were presented to participants on an Apple Macbook laptop using the Psyscope software package (see Experiment 1). Stimuli consisted of twenty-four nonsense syllables (see Appendix 2) arbitrarily assigned to their roles as samples, comparisons, and FAST test stimuli (see below). These will be referred to hereafter using alphanumerics.

Ethics

All participants were presented with and signed a consent form before proceeding to first phase of the experiment (See Appendix 1). Participants were told informally that performance on the task would not allow the researcher to make any individual psychological assessments but may allow for group patterns to be identified. All participants were informed of the true nature of the study after participation and were given the opportunity to withdraw their data at that stage.

General Experimental Procedure

The experiment consisted of two phases. Phase one (A-B training) consisted of a linear matching-to-sample protocol to establish two simple stimulus relations between two pairs of nonsense syllables (A1-B1-C1 and A2-B2-C2). Phase two (FAST) consisted of three runs (exposures) of a pair of FAST test blocks, separated by a baseline block, and with an additional baseline block presented at the end of the entire procedure. Multiple runs of the FAST were employed so that the stability of participant performance across time could be considered.

Phase 1: Matching-To-Sample Training. In this first Phase, participants were exposed to four matching-to-sample training tasks, each designed to establish two simple stimulus relations. The relations trained were; choose B1, not B2 when A1 is present; choose B2, not B1 when A2 is present, choose A1 not A2, when B1 is present, and choose A2 not A1 when B2 is present (see Figure 1). The purpose of this phase was to establish two laboratory controlled stimulus relations that could be employed to assess the utility of the FAST in determining the existence and strength of known stimulus relations.

Participants were presented with the following instructions at the onset of Phase 1.

In a moment some words will appear on this screen. Your task is to look at the word at the top of the screen and choose one of the two words at the bottom of the screen by "clicking on it" using the computer mouse and cursor. During this stage the computer will provide you with feedback on your performance. You should try to get as many answers correct as possible. If you have any questions please ask them now. When you are ready please click the mouse button.

All trials were presented on the computer screen against a white background. A trial began with the presentation of the sample stimulus at the top centre of the screen in black 24-point font. One second later, the two comparison stimuli were displayed in black 24-point font in the bottom left and right corners of the screen (see Figure 1). The positions of the comparison stimuli were counterbalanced across trials. All stimuli remained onscreen until the participant responded by clicking on one of the comparison stimuli using the computer mouse pad and cursor. Immediately upon a response, the screen cleared and corrective feedback ("Correct" or "Wrong" in red 48 point font) appeared in the centre of the screen for 1.5s.

Trials were presented in blocks of 32 (i.e., each of the four trials presented 8 times). Participants were required to complete successive training blocks (up to a maximum of four cycles) until a criterion of 31 or 32 correct responses in a single 32-trial block was met (i.e., minimum 96.9% correct).

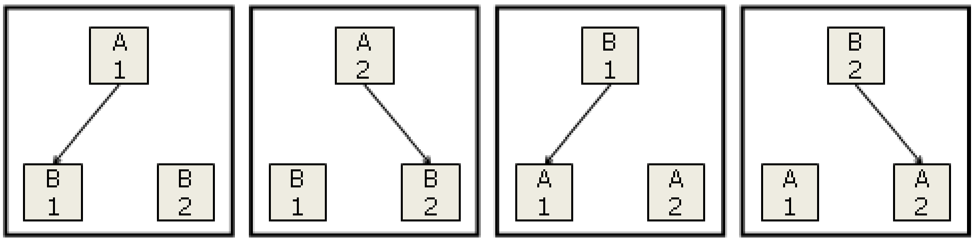


Figure 1. The stimulus relations trained during Phase 1. Solid lines indicate reinforced matching responses.

Phase 2: The Function Acquisition Speed Test

Phase 3 consisted of three consecutive Function Acquisition Speed Tests. The first of these was the critical test. The subsequent tests were administered to consider the robustness of FAST effects across repeated immediate exposures. The reader is reminded that to thus point a typical FAST presentation consists of a baseline block, two test blocks (consistent and inconsistent), and a further baseline block. In this experiment, participants were exposed to three sets of test blocks, each preceded by a baseline block, with a final baseline block at the end, thus producing the sequence; Baseline 1, FAST Test Blocks 1, Baseline 2, FAST Test Blocks 2, Baseline 3, FAST Test Blocks 3, Baseline 4 (see Figure 6, below).

General FAST Blocks Procedure

Stimulus presentation, feedback, trial order, response windows, and maximum trial limit were identical to the *General FAST Block Procedure* used in Experiment 3. Figure 2 below describes the block order and required responses for the FAST procedure.

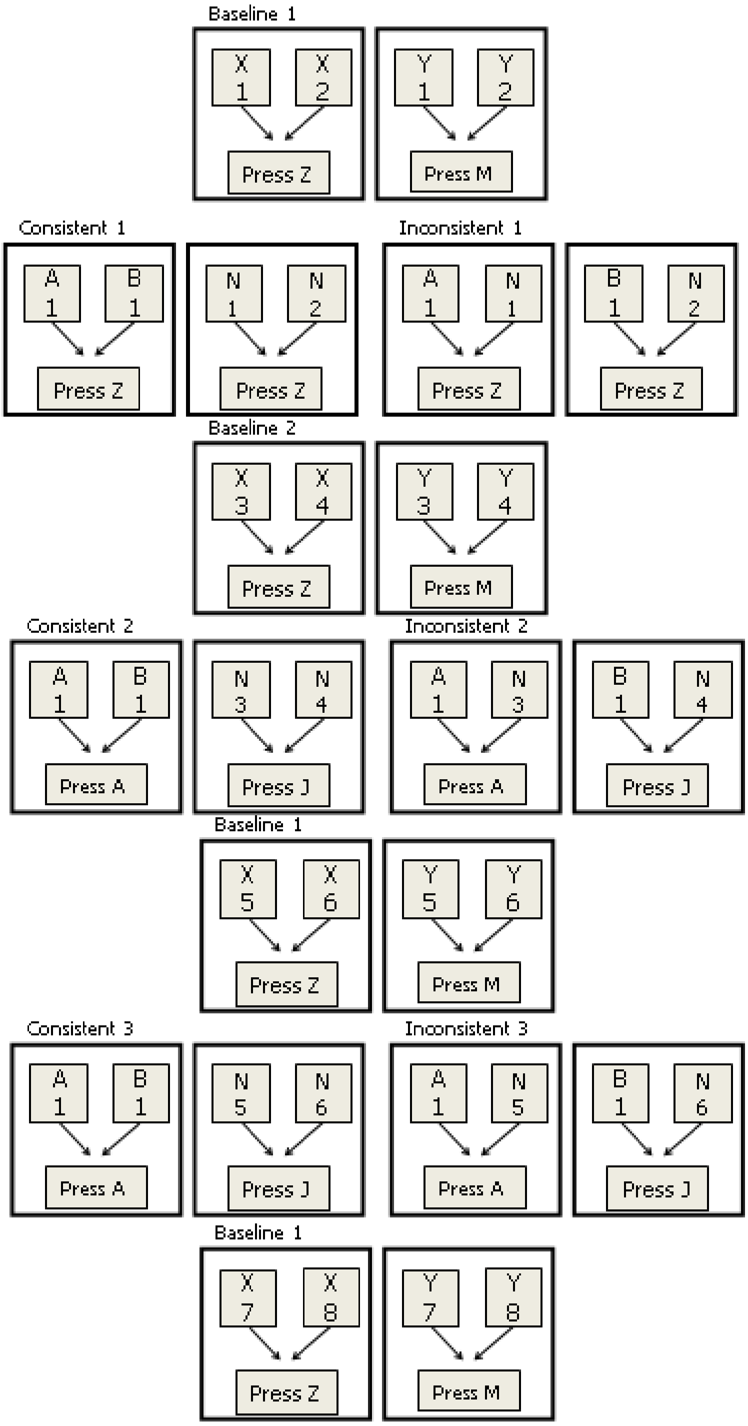


Figure 2. A schematic of the block order and various response functions assigned to stimuli during baseline and FAST blocks. The panels, read from top to bottom, represent each FAST block presented in sequence.

Baseline Blocks

The procedure used for the Baseline Blocks was identical to that used in Experiment 3.

Test Blocks

The test blocks utilized the A1 and B1 stimuli from the Equivalence Training and Testing Phases, and two further novel nonsense syllables, N1 and N2. One of these blocks (consistent) established two functional response classes (A1-B1 and N1-N2) that were consistent with the derived relations predicted given Phase 1. The other block established two functional response classes that were inconsistent with the relations trained in Phase 1 (A1-N1 and B1-N2). The order of the consistent and inconsistent blocks was randomized across participants.

The stimulus presentation procedure used in the test blocks was identical to that used in the Baseline Blocks – only the stimuli used differed between blocks. The test blocks also utilized different response keys (*z* and m) to prevent any conflicting response histories across baseline and test blocks.

The instructions presented to participants at the onset of the FAST Baseline Blocks and Test Blocks were identical to those presented in Experiment 3.

To summarize, all participants were exposed to a series of blocks in the following sequence; Baseline 1, FAST 1, Baseline 2, FAST 2, Baseline 3, FAST 3, and Baseline 4. After the completion of each block, the instructions page for the next block appeared, allowing the participant to begin the next block whenever they were ready without interruption from the experimenter. Upon completion of the final baseline block, a page appeared thanking the participant for their participation and instructing them to contact the experimenter. The participants were split into four experimental groups. For each group, the response window allowed for responses was 1000, 2000, 3000, or 4000ms, respectively.

Results

Baseline Blocks

Participants completed four baselines over the course of the experiment. For Group 1 (1s response window), the mean baseline trials to criterion were 44.071 (SD = 22.797). For Group 2 (2s Response Window), the mean baseline was 23.985 (SD = 7.781). For Group 3 (3s Response Window) the mean baseline was 34.1944 (SD = 20.675). For Group 4 (4s response window) the mean baseline was 30.775 (SD = 16.645). A one-way ANOVA showed no significant differences in mean baseline between the experimental groups. However, it is worth noting that the Standard Deviation for Group 2 was a great deal smaller than the other experimental groups, indicating a more stable baseline performance within that group as compared to the other response windows.

A series of paired sample t-tests showed no significant practice effects across the four baselines. This outcome served as a first indication that response windows have inconsistent effects on task performance because baseline performances were also expected to differ across the four Groups. While Group 1 did indeed display the largest trial requirement, as expected, the remaining three groups showed varying trial requirements that are difficult to relate to response window duration,

Function Acquisition Speed Test

Group 1 (1s Response Window). Seven participants completed the FAST procedure with a 1s response window. In the first run of the FAST, four of the seven participants showed a FAST effect in the expected direction (i.e. trials to criterion on consistent block < inconsistent block). The mean difference between inconsistent and consistent blocks in the expected direction was 15 (SD =8.66). The mean difference in the unexpected direction was 20. 428 (SD = 20.428). The overall mean difference between inconsistent and consistent blocks was 1.571 (SD = 23.901).

Using the same formula as in previous experiments, an SOR index was calculated. The mean SOR was 0.4622 (SD = 6.218). A Wilcoxon signed ranks test showed that this was not significantly different from zero (z = .507, p = .612), although it must be remembered that the sample was very small.

On the second run of the FAST, four out of seven participants in Group 1 showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 26 (SD = 30.276.66). The mean difference in the unexpected direction was 9.333(SD = 10.115). The overall mean difference between inconsistent and consistent blocks was 10.85 (SD = 29.14). The mean SoR on the second run was 3.5935 (SD = 9.687). However, this was not significantly different from zero, according to a Wilcoxon signed ranks test. (z = .507, p=.612)

On the third run of the FAST, five out of seven participants in Group 1 showed a FAST effect in the expected direction. One participant showed a negative effect, while one participant showed no response requirement difference across the FAST blocks. The mean difference between inconsistent and consistent blocks in the expected direction was 20.6 (SD = 5.458). The difference in the unexpected direction was -14. The overall mean difference between inconsistent and consistent blocks was 12.71 (SD = 14.75). The mean SoR on the third and final run was 3.689 (SD = 4.185). A Wilcoxon signed ranks test showed that this difference was significantly difference from zero (z = 2.023, p = .046).

Group 2 (2s Response Window). Six participants completed the FAST procedure with a 2s response window. In the first run of the FAST, four of the six participants showed a FAST effect in the expected direction. The mean difference in the expected direction was 15 (SD = 20.215). The mean difference in the unexpected direction was -20.5. (SD = 16.263). The overall mean difference between inconsistent and consistent blocks was 3.166 (SD =

25.182). The mean SOR was 0.4904 (SD = 8.045). A Wilcoxon signed ranks test showed that this was not significantly different from zero (z = .524, p = .60).

On the second run of the FAST, two out of six participants showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 11.5 (SD = 4.949). The mean difference in the unexpected direction was 20.5(SD = 22.883). The overall mean difference between inconsistent and consistent blocks was -9.833 (SD = 24.33). The mean SoR on the second run was -3.758 (SD = 5.274). However, this was not significantly different from zero, according to a Wilcoxon signed ranks test (z = -0.943, p = 0.345).

On the third run of the FAST, five out of six participants showed a FAST effect in the expected direction. One participant showed no difference. The mean difference between inconsistent and consistent blocks in the expected direction was 18.6 (SD = 9.939). The overall mean difference between inconsistent and consistent blocks was 15.5 (SD = 11.69). The mean SoR on the third and final run was 5.011 (SD = 4.026). A Wilcoxon signed ranks test showed that this difference was significantly difference from zero (z = 2.023, p = .043).

Group 3 (3s Response Window). Nine participants completed the FAST procedure with a 3s response window. In the first run of the FAST, four of the nine participants showed a FAST effect in the expected direction. One participant showed no difference. The mean difference in the expected direction was 2.5 (SD = 2.38). The mean difference in the unexpected direction was 24.5 (SD = 28.407). The overall mean difference between inconsistent and consistent blocks was -9.778 (SD = 22.371). The mean SOR was -3.0454 (SD = 6.925). A Wilcoxon signed ranks test showed that this was not significantly different from zero (z = -.840, p = .401).

On the second run of the FAST, only three out of nine participants in Group 3 showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 5.667 (SD = 2.081). The mean difference in the unexpected direction was 18.333(SD = 22.482). The overall mean difference between inconsistent and consistent blocks was -10.333 (SD = 21.471). The mean SoR on the second run was -2.758 (SD = 5.275). However, this was not significantly different from zero, according to a Wilcoxon signed ranks test (z = -1.481, p = 0.139).

On the third run of the FAST, eight out of nine participants in Group 3 showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 21.375 (SD = 19.711). The difference in the unexpected direction was 12. The overall mean difference between inconsistent and consistent blocks was 17.667 (SD = 21.534). The mean SoR on the third and final run was 5.577 (SD = 6.681). A Wilcoxon signed ranks test showed that this difference was significantly different from zero (z = 2.310, p = .021).

Group 4 (4s Response Window). Ten participants completed the FAST procedure with a 3s response window. In the first run of the FAST, five of the ten participants showed a FAST effect in the expected direction. The mean difference in the expected direction was -4.4 (SD =12.234). The mean difference in the unexpected direction was 16.333. (SD = 14.556). The overall mean difference between inconsistent and consistent blocks was -4.4 (SD = 20.122). The mean SOR was -1.613 (SD = 5.753). A Wilcoxon signed ranks test showed that this was not significantly different from zero (z = -.764, p = .445).

On the second run of the FAST, two out of ten participants showed a FAST effect in the expected direction. One participant showed no difference. The mean difference between inconsistent and consistent blocks in the expected direction was 39 (SD = 25.455). The mean difference in the unexpected direction was -20.428 (SD = 19.544). The overall mean difference between inconsistent and consistent blocks was -6.5 (SD = 30.695). The mean SoR on the second run was -1.803 (SD = 8.838). However, this was not significantly different from zero, according to a Wilcoxon signed ranks test (z = -1.125, p = 0.260).

On the third run of the FAST, six out of ten participants showed a FAST effect in the expected direction. One participant showed no difference. The mean difference between inconsistent and consistent blocks in the expected direction was 15.667 (SD = 14.236). The mean difference in the unexpected direction was -16 (SD = 22.538). The overall mean difference between inconsistent and consistent blocks was 4.6 (SD = 21.235). The mean SoR on the third and final run was 1.581 (SD = 5.541). A Wilcoxon signed ranks test showed that this difference was significantly different from zero (z = 1.007, p = .314).

Summary

In the first run of the FAST test blocks, a majority of participants showed effects in the expected direction in the 1s and 2s groups. For the 3s and 4s groups, the opposite was true. In the second run, many participants showed a reversal of their scores on the first FAST run. Only the 1s group showed a majority of effects in accordance with the pre-experimental A-B training. However, in none of the groups were any of the SoR scores significantly different from zero for any of the first two FAST runs. This was most likely due to small group size.

In the third and final run, the 1s, 2s, and 3s groups showed a majority completing the consistent blocks more quickly than the inconsistent blocks. This seems to indicate that test effects stabilise across the first to the second and third exposure. Further, these results were shown to be statistically significant from zero when the SoRs were tested using Wilcoxon signed ranks tests. The 4s group showed a similar pattern, but the results were not statistically significant. Thus, a broad conclusion at this point is that, while statistical analyses are not particularly meaningful with such small sample sizes, there is a general trend towards larger

FAST effects with shorter response windows. Regardless of the significance or otherwise of FAST effects observed with these small samples, there is a definite and marked trend towards a proliferation of effects in the expected direction when response windows are large compared to small.

Excluded Participants

Twelve participants were excluded from the above data analysis on the basis of their failure to complete at least one FAST block before the 100 trial limit was reached. It is worth noting, however, that of these twelve participants, 4 participants were in Group 1, 2 participants were in Group 2, 3 were in Group 3, and 2 were in Group 4. The largest number of participants failing to complete a block within 100 trials was found amongst those who were required to complete the FAST trials within 1000ms. This suggests that a 1s response window may be prohibitively difficult in comparison to other response windows. Thus, while shorter response windows tend to lead to better control over test effects, an overly stringent response window may in fact lead to the loss of too much data through the complete failure of participants to perform the test.

Between Group Effects.

The SoR scores from each of the four groups on each run of the FAST were compared using a one-way ANOVA. The first run of the FAST lead to no significant differences in FAST effect sizes across the four groups ($F = 1.035$, $p = .14$). The same was true for the second ($F = 2.725$, $p = .057$) and third ($F = 1.702$, $p = .183$) runs of the FAST. In effect, an analysis of variance has failed to indicate a significant change in FAST effect sizes that can be associated with response window size.

Discussion

The current experiment aimed to compare the effects of different response windows on FAST performances. The first three groups (1s, 2s, 3s) showed a similar pattern of SoRs across multiple runs, showing a significant FAST effect at the group level in the third and final FAST blocks. The fourth group (4s) showed no significant FAST effect on any run. This suggests that greater time pressure is more fruitful for implicit testing methodologies. Indeed, IRAP research (Barnes-Holmes et al., 2010) has found that a reduction from a 3000ms to a 2000ms response window increased IRAP effects. While the IRAP is a different procedure harnessing the same basic behavioural process, it comes as little surprise that increased time pressure, within boundaries, would increase the ability of the ability of an implicit test to detect a pre-experimentally trained history. While, no significant statistical difference was discovered between 1s, 2s, and 3s response window FAST in terms of the size of SoRs, the general trend towards more prediction and influence over behaviour is marked. The fact that it required multiple runs of the FAST for a significant effect to occur is suggestive of the need for practice blocks to be included in the FAST procedure. The role of practice will be considered in the following experiment.

There were some differences in the patterns of results displayed by each group. The mean number of trials to criterion across all blocks was 39.74 for Group 1, 21.653 for Group 2, 26.833 for Group 3, and 27.68 for Group 4. Further, the largest group of participants eliminated for reaching the trial limit came from Group 1. Participants who completed the FAST with a 1s response window required more trials to complete any block, regardless of contingencies, and were more likely to fail to complete a block within 100 trials. This suggests that a 1s response window may be too small for use in future studies. While the raw trial requirement is not the primary datum in a FAST, it is relevant here insofar as we are also concerned with participants failing to complete a FAST procedure within the required number of trials.

The 2s and 3s groups both demonstrated an unexpected reversal in group FAST effects on the second run of the FAST, before stabilizing on the third run to produce a significant FAST effect. This surprising pattern of results was not present, in the 1s response window condition. The responses of participants in that condition were more consistent across multiple exposures to the FAST. This may support the further use a 1s response window criterion, but the increased difficulty it creates for participants mitigates against its use. One of the aims of a test such as the FAST is for it to be administered by all and any individuals and not to be so complex that it requires a certain level of pre-established ability in such tasks for the test format to be of any use.

It is difficult to be certain of the source of the sudden (though statistically non-significant) reversal of FAST effects observed in the above data. In Experiment 2, we observed similarly unstable FAST performances across multiple runs. In contrast to Experiment 2, the current participants did not display a significant FAST effect until the final run, whereas in Experiment 2, the significant FAST effect emerged on the second FAST run; the point at which we observed a reversal of FAST effect direction in the current experiment.

As was noted in Experiments 2 and 3, it is possible that a lack of robustness in the training procedures might account for some of the observed instability within participant performers. The current experiments were conducted before Experiment 4, but the results of this experiment support this conclusion and go some way towards explaining the pattern of results.

One likely source of the patterns observed in these experiments is the manner in which block order was determined. At the onset of each pair of FAST test blocks, the order (consistent first or inconsistent first) was randomized. In experiments with only one FAST run, order was not found to be a factor. As such, it was not expected that order would be problematic if it continued to be randomized. However, this issue emerged in each experiment described in the current chapter. All three of these experiments were run simultaneously, and as such the researcher was unable to address the issue until data from all experiments was analysed. To avoid repetition, this issue will be discussed in the General Discussion of this chapter.

In conclusion, though the results were not perfectly clear, the trends observed in the current experiment suggest that a lower response window should be used where possible. However, a 1s response window may result in too many participants failing to complete the procedure due to increased difficulty. The trade off between lost participants and an increased amount of time pressure would be best resolved with implementing a 2000ms response window in future experiments. Future research should consider the issue of response windows, and examine a narrow range of low response windows (between 1000 and 2000ms) in light of the procedural refinements introduced in Experiment 4.

Experiment 7

Exploring the Effects of Practice

The FAST operates by requiring participants to learn to assign a response (either press left or press right) to each of four stimuli. Participants must learn to emit these responses both quickly (i.e., within a response window) and accurately (with 1 mistake or fewer within a succession of ten trials) based only on feedback presented after each trial. Baseline blocks utilise unrelated stimuli to establish a participants' normal ability to complete a task of this sort in the absence of any pre-experimental relations. The FAST procedure is an unusual task for most participants, who are not likely to have had experience with similar procedures.

In previous experiments, multiple baselines have been utilized, so as to produce a mean baseline score which can be used in calculating the SoR Index. In each of these experiments, a general decrease in trials to criterion in the baselines has been observed as time progresses. Unsurprisingly, as participants complete more FAST tasks, they become more proficient at identifying the correct responses, perhaps developing strategies to receive useful feedback as quickly as possible when presented with novel stimuli. For instance, responding "left" four times at the start of a baseline block allows a participant to glean all relevant information in the minimum time for all possible arrangements of trial order. Depending on trial order, it is possible to gain all necessary information to complete a FAST block in two trials. As such, when a participant is attentive to multiple exposures of the FAST procedure, overall performance can be highly accelerated.

On the other hand, the FAST procedure requires rapid responding, and proceeds through trials without necessarily receiving any input from the participant. If the Baseline

Block is presented immediately, confusion about rules, or simply the rapid pace of the test, could conspire to produce a Baseline score that is higher than would be expected. This serves to reduce the size of SoR indices in subsequent data analyses.

The current experiment seeks to explore the effect of providing a number of baseline-style practice blocks to participants undergoing a FAST procedure. The aim of this experiment is to explore the rate at which practice block performance reaches a stable plateau over multiple exposures, and to investigate when, if at all, performances approach a floor effect wherein participants complete all subsequent FAST blocks in a very small amount of trials.

Method

Participants

Twenty three volunteers participated in the study, 11 of which were male and 12 of which were female. The participants' ages ranged from 18-47 (Mean = 21.52, SD = 1.81). Of these, 6 participants were excluded from initial data analysis due to failure to complete one or more test blocks within the 100 trial limit.

Apparatus and Stimuli

All phases of the experiment were presented to participants on an Apple Macbook laptop computer with a 13-inch monitor (1024 x 768 pixel resolution). Stimulus presentations were controlled by the Psyscope software package (Cohen, MacWhinney, Flatt & Provost, 1993) which also recorded all responses. Stimuli consisted of 40 nonsense syllables randomly assigned to their roles as samples, comparisons, and FAST test stimuli (see Appendix 2). These will be referred to in future using alphanumerics.

Ethics

All participants were presented with and signed a consent form before proceeding to first phase of the experiment (See Appendix 1). Participants were told informally that performance on the task would not allow the researcher to make any individual psychological assessments but may allow for group patterns to be identified. All participants were informed of the true nature of the study after participation and were given the opportunity to withdraw their data at that stage.

General Experimental Procedure

The experiment consisted of three phases. Phase 1 (A-B Training) consisted of a matching to sample protocol to establish two simple stimulus relations (A1-B1- and A2-B2). All participants received identical training in stimulus-stimulus matching.

Phase 2 consisted of a number of practice blocks, varied by experimental group. Group 1 received 3 practice blocks, Group 2 received 5 practice blocks, and Group 3 received 7. These practice blocks used the same format as FAST Baseline blocks, with participants being required to form two functional response classes from four novel, unrelated nonsense syllables.

Phase 3 consisted of the Function Acquisition Speed Test, and it probed for the A1-B1 relation trained in phase 1. Each of the three FAST exposures consisted of two test blocks (consistent and inconsistent), with four additional baseline blocks interspersed between them (Baseline 1, Test Blocks 1, Baseline 2, Test Blocks 2, Baseline 3, Test Blocks 3). The first pair of FAST blocks were the main blocks of interest, but the additional FAST exposures were intended to identify varying effects of practice on FAST test outcomes on both first and subsequent exposures.

All phases were presented consecutively by the computer software, which also controlled the delivery of instructions at the beginning of each phase. Participants sat comfortably at a standard computer desk and viewed the screen at a distance of approximately 60-70cm and at eye level.

Phase 1: A-B Training

This phase was identical to Phase 1 in Experiment 6. Figure 3 below details the relations trained.

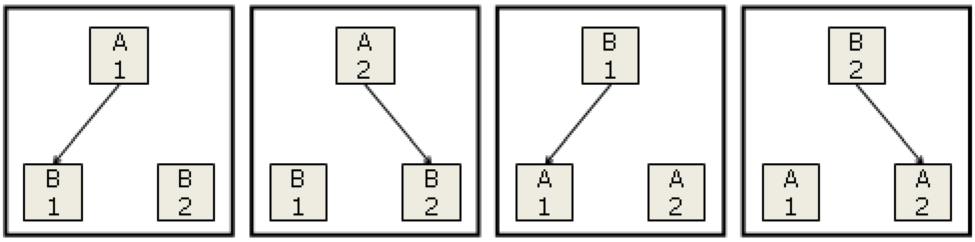


Figure 3. The stimulus relations trained during Phase 1. Solid lines indicate reinforced matching responses.

Phase 2: FAST Practice Blocks

The number of practice blocks administered to participants varied by experimental group. Group 1 was exposed to 3 practice blocks, Group 2 was exposed to 5, and Group 3 to 7. The purpose of this experimental manipulation was to determine the optimal number of practice blocks to stabilize participants' performances and minimize variation in test effects across individuals.

Participants received the following instructions at the start of each practice block.

> *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.*

All trials were presented on the computer screen with a white background. A trial began with the presentation of one of four nonsense syllable stimuli (e.g., for practice block 1; P1, P2, Q1, Q2) in the centre of the screen in 48-point black font. Each of the four stimuli was presented in a quasi-random order in blocks of four trials (i.e., no more than two consecutive exposures to any one stimulus was possible).

A strict response window was enforced on participants' responses. Stimuli remained on screen for a 3000ms, after which, if a response had not been emitted, the screen would clear, the trial would end, and an "incorrect" response would be recorded by the computer.

Immediately upon the production of a response, corrective feedback was presented (i.e. either "Correct" or "Wrong" presented in red 48-point font in the centre of the screen for 1.5s). If no response was emitted within the 3s response window, an incorrect response was recorded but no feedback was provided. In that case, the subsequent trial began immediately upon the end of the 3s response window. Participants were exposed to trials until a criterion of 9/10 correct in a row was reached. That is to say, upon the emission of a first correct response, the program began to count to ten. Participants were allowed to make one error. On the second error, the count reset to zero. If the participant reached 10 responses with less than 2 errors (i.e., 9/10 correct) then the block ended.

The practice blocks involved novel and unique stimuli. Practice block 1 employed P1, P2, Q1, and Q2 as stimuli, while Practice bloc 2 employed P3, P4, Q3, and Q4 as stimuli, and so on. These phases attempted to establish the following functional stimulus relations, respectively; P1-Q1, P2-Q2 and P3-Q3, P4-Q4. That is, when P1 or Q1 appeared on screen, participants were delivered with reinforcement for pressing m, and when P2 or Q2 appeared on screen, pressing z was reinforced.

Phase 3: The Function Acquisition Speed Test

The number of blocks, order of presentation, instructions, and response criteria were identical to those used in Experiment 6. The response window for all participants was held constant at 3000ms for all trials due to the fact that this experiment was run in parallel with Experiment 6 and the effects of response windows on FAST effects was not yet known. Figure 4 below describes the block order and required responses for the FAST.

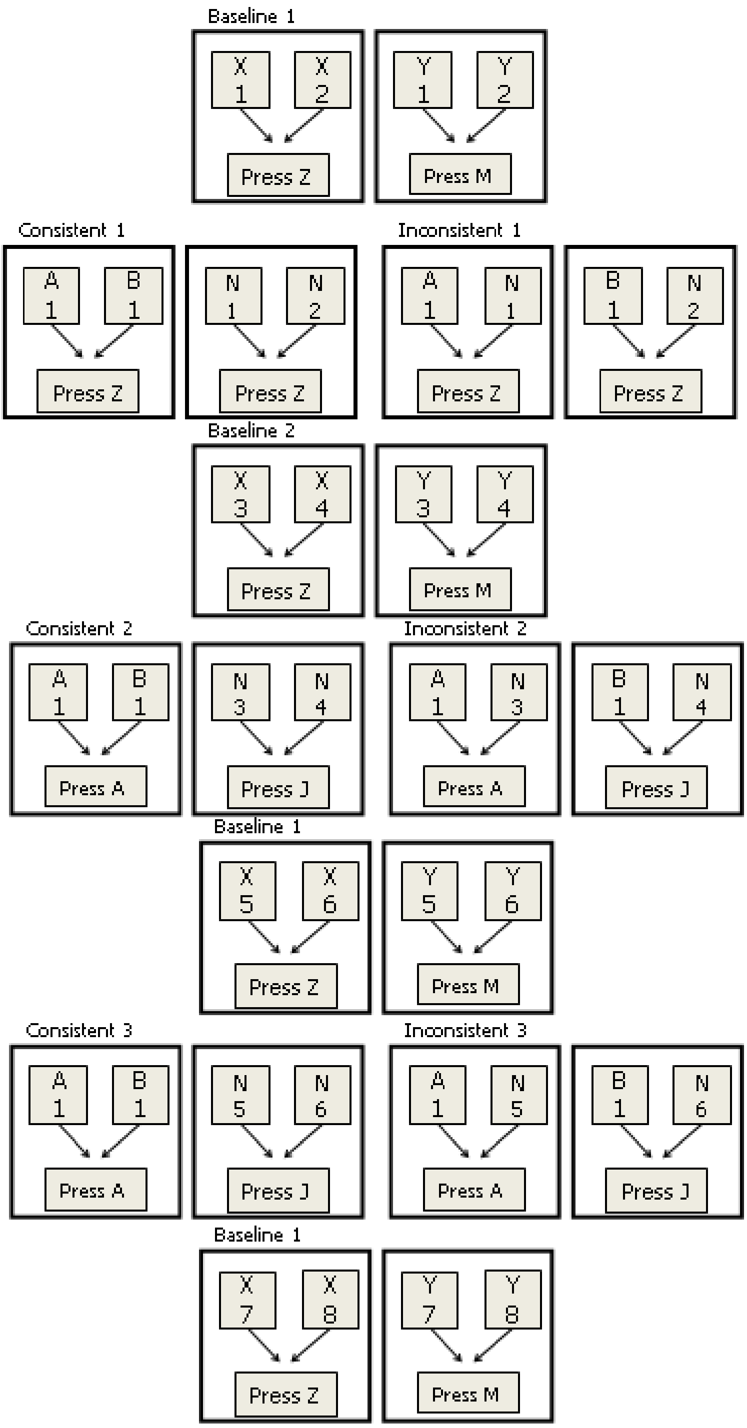


Figure 4. A schematic of the block order and various response functions assigned to stimuli during baseline and FAST blocks. The panels, read from top to bottom, represent each FAST block presented in sequence.

Results

Practice Blocks

Group 1 (3 Practice Blocks). Six participants completed three practice blocks. The mean trials-to-criterion for the first practice block was 47.50 (SD = 27.157). For the second practice block, the mean was 62.33 (SD = 29.73), and for the third, the mean was 73 (SD = 35.094). Surprisingly, the trial requirement climbed, rather than fell, from block to block. However, a related samples Friedman's Two Way Analysis of Variance by Ranks showed that this difference was non-significant (z = 4.261, p = .119).

Group 2 (5 Practice Blocks). Four participants completed five practice blocks. The mean trials-to-criterion for the first practice block was 29.75 (SD = 11.206). For the second practice block, the mean was 48.75 (SD = 44.969), for the third, the mean was 42.75 (SD = 20.073). For the fourth practice block, the mean was 73.25 (SD = 36.363), and for the fifth the mean was 57.25 (SD = 29.273). Once again trial requirements generally rose rather than fell across practice blocks, although there was a drop from block two to block 3. A related samples Friedman's Two Way Analysis of Variance by Ranks showed that this difference in trial requirements across runs was non-significant (z = 6.773, p = .148).

Group 3 (7 Practice Blocks). Seven participants completed seven practice blocks. The mean trials-to-criterion for the first practice block was 25.43 (SD = 11.088). For the second practice block, the mean was 31.14 (SD = 27.046) and for the third, the mean was 44.43 (SD = 33.970). For the fourth practice block, the mean was 73.25 (SD = 36.363), and for the fifth the mean was 47.86 (SD = 41.261). The mean for the sixth practice block was 29.29 (SD = 16.730) and for the seventh, 30 (SD = 22.635). A similar pattern was once again observed in which trial requirements actually rose rather than fell across runs, although the trial requirements appears to plateau somewhat and eventually fall after run 5. A related samples Friedman's Two Way Analysis of Variance by Ranks showed that this difference was non-significant (z = 3.332, p = .766).

Overall, no statistical difference was found between participants' performance across consecutive practice blocks. However, participants' trial requirements to block completion tended to increase as practice blocks continued until a peak at block 4, after which a drop was observed from block 4 to five. Group 3 continued to show this drop, with a small increase from Blocks 6 to 7, suggesting a possible plateau. Figure 5 below graphs the mean trials to criterion for each group across their respective practice blocks.

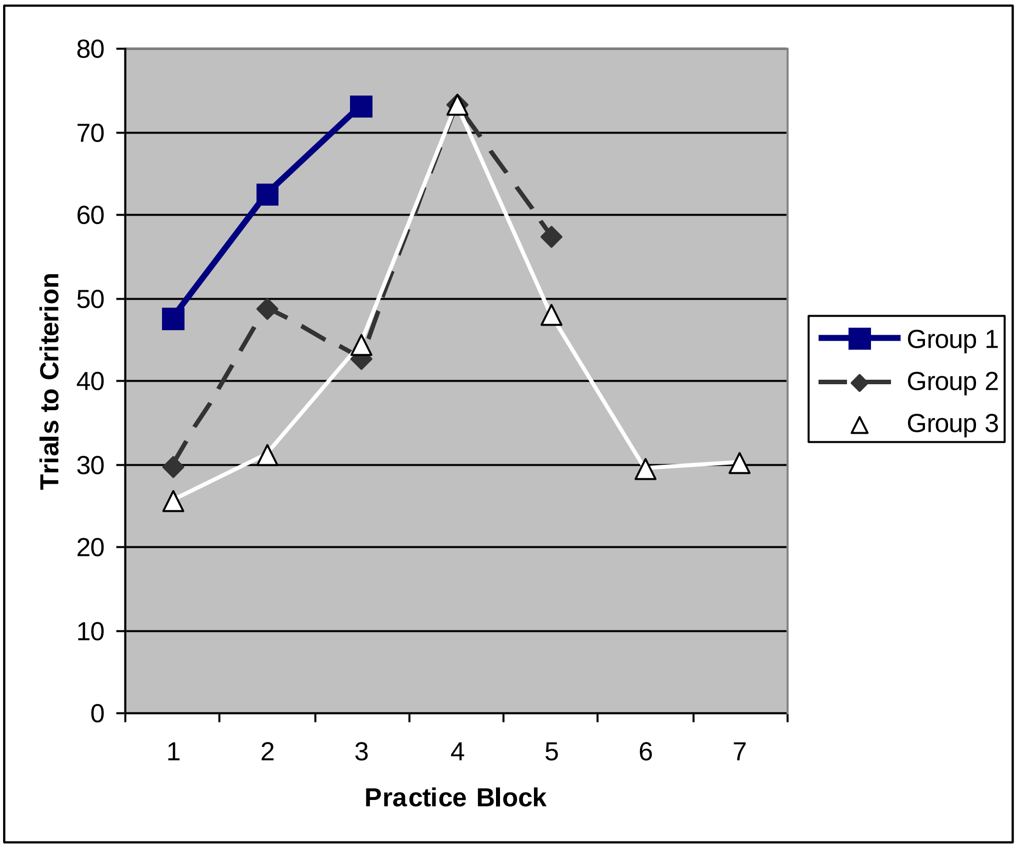


Figure 5: Mean trials to Criterion on each practice block for each experimental group.

Baseline Blocks

Participants completed four baseline blocks over the course of the experiment. For Group 1 (3 Practice Blocks), the mean baseline was 32.791 (SD = 16.493). For Group 2 (5 Practice Blocks), the mean baseline was 32.75 (SD = 14.29). For Group 3 (7 Practice Blocks) the mean baseline was 27.464 (SD = 17.333). A one-way ANOVA showed no significant differences in mean baselines between the experimental groups (F = .215, p = .809). Interestingly, the temporal spacing of the baselines across FAST runs, and the fact that they came subsequent to the practice period, seems to have resulted in quite stable baseline performances across groups, irrespective of the amount of practice provided.

Function Acquisition Speed Test

Group 1 (3 Practice Blocks). In the first run of the FAST, four of the six participants showed a FAST effect in the expected direction (i.e. trials to criterion on consistent block < inconsistent block). The mean difference between inconsistent and consistent blocks in the expected direction was 12.75 (SD =12.842). The mean difference in the unexpected direction was 34.5 (SD = 10.607). The overall mean difference between inconsistent and consistent blocks was -3 (SD = 26.773). The mean SOR was -0.81 (SD = 7.741). A Wilcoxon signed ranks test showed that this was not significantly different from zero (z = .105, p = .917).

On the second run of the FAST, two out of six participants showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 17 (SD = 7.616). The mean difference in the unexpected direction was -8(SD = 7.616). The overall mean difference between inconsistent and consistent blocks was 0.333 (SD = 15.069). The mean SoR on the second run was -0.386 (SD = 4.328). However, this was not significantly different from zero, according to a Wilcoxon signed ranks test (z= -0.314, p = .753).

On the third run of the FAST, five out of six participants in Group 1 showed a FAST effect in the expected direction. One participant showed a negative effect. The mean difference between inconsistent and consistent blocks in the expected direction was 34.8 (SD = 30.817). The difference in the unexpected direction was -12. The overall mean difference between inconsistent and consistent blocks was 27 (SD = 33.538). The mean SoR on the third and final run was 8.1189 (SD = 9.096). A Wilcoxon signed ranks test showed that this difference from zero approached significance (z = 1.782, p = .075).

Group 2 (5 Practice Blocks). In the first run of the FAST, three of the four participants showed a FAST effect in the expected direction. The mean difference in the expected direction was 27.667 (SD = 18.556). The difference in the unexpected direction was -5. The overall mean difference between inconsistent and consistent blocks was 19.5 (SD = 22.278). The mean SoR was 5.855 (SD = 7.376). A Wilcoxon signed ranks test showed that this was not significantly different from zero (z = 1.461 p = .144).

On the second run of the FAST, three out of four participants in Group 2 showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 23 (SD = 35.071). The difference in the unexpected direction was -5. The overall mean difference between inconsistent and consistent blocks was 32.333 (SD = 36.364). The mean SoR on the second run was 6.227 (SD = 9.255). However, this was not significantly different from zero, according to a Wilcoxon signed ranks test (z = 1.461, p = 0.144).

On the third run of the FAST, two out of four participants in Group 2 showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 13 (SD =1.414). The mean difference in the unexpected direction was -33.5 (SD = 43.134). The overall mean difference between inconsistent and consistent blocks was -10.25 (SD = 36.627). The mean SoR on the third and final run was -2.533 (SD = 9.771). A Wilcoxon signed ranks test showed no significant difference from zero (z = 0, p = 1).

Group 3 (7 Practice Blocks). In the first run of the FAST, five of the seven participants showed a FAST effect in the expected direction. One participant showed no difference. The mean difference in the expected direction was 10.2 (SD = 7.19). The difference in the unexpected direction was -13.The overall mean difference between inconsistent and consistent blocks was 5.429 (SD = 10.721). The mean SoR was 1.582 (SD = 6.93.935). A Wilcoxon signed ranks test showed that this was not significantly different from zero (z = 1.153, p = .249).

On the second run of the FAST, three out of seven participants showed a FAST effect in the expected direction. Two participants showed no difference. The mean difference between inconsistent and consistent blocks in the expected direction was 11.667 (SD = 9.866). The mean difference in the unexpected direction was -18.5 (SD = 24.748). The overall mean difference between inconsistent and consistent blocks was -0.286 (SD = 17.792). The mean SoR on the second run was -0.176 (SD = 3.936). However, this was not significantly different from zero, according to a Wilcoxon signed ranks test (z = .405, p = 0.686).

On the third run of the FAST, five out of seven participants in Group 3 showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 10.4 (SD = 10.114). The mean difference in the unexpected direction was -14.5 (SD = 6.369). The overall mean difference between inconsistent and consistent blocks was 3.286 (SD = 14.919). The mean SoR on the third and final run was 0.923 (SD = 4.913). A Wilcoxon signed ranks test showed that this difference was not significantly difference from zero (z = .676, p = .499).

Summary

No groups showed a statistically significant group level FAST effect in any of the three FAST runs. However, a majority of participants across groups showed a positive SoR in FAST runs 1 and 3, with just under half showing the effect in FAST run 2. The lack of statistical significance can be attributed to low sample sizes but possibly also to a paradoxical disrupting effect of practice on test effects. That is, it seems that the very provision of practice consisting of at least three separate blocks, served to interfere with learning and the disruptive effects seem to have increased across practice blocks (see Discussion).

Excluded Participants

Six participants were excluded from the above data analysis on the basis of their failing to complete a FAST block before the 100 trial limit was reached. Of these twelve participants, 1 participant was in Group 1, 4 participants were in Group 2, and 1 was in Group 3. The largest number of participants failing to complete a block within 100 trials was found amongst those who were required to complete five practice blocks.

Between Group Effects.

The SoR scores from each of the three groups on each run of the FAST were compared using a one-way ANOVA. The first run of the FAST showed no significant differences ($F = 1.347$, $p = .292$). The same was true for the second ($F = 1.808$, $p = .2$) and third ($F = 2.55$, $p = .114$) runs of the FAST.

Discussion

Participants did not show a linear decrease in trials-to-criterion across multiple Practice Blocks. While the differences were non-significant, there tended to be an increase in trials-to-criterion as participants completed more practice blocks, until the fifth practice block for Groups 2 and 3. This is somewhat surprising, as a general decrease would be expected as participants become more familiar with the testing procedure.

However, consider the pattern that emerged in the current experiment – a sudden drop in requirements on the fifth practice block. In prior experiments which used upwards of five blocks (counting the test blocks), a statistically significant practice effect was observed. While in each of these experiments, the FAST effect emerged across the groups (with the exception of Experiment 2), there was a noticeable drop in the second baseline score. The second (and any subsequent) baseline uses novel nonsense stimuli that have not been used in any previous block of the FAST. As such, and as intended, the participants' responses on a baseline block are not under the control of previous experiences with the stimuli being used. As such, Baseline Blocks tend towards having higher trial to criterion requirements than Test Blocks.

However, when a second (or any subsequent) baseline is presented, the participant has already completed at least three FAST blocks. Putting aside for the moment the fact that two of these will have been Test Blocks, a participant has enough experience with the format itself that they may have been able to generate an effective strategy which generalizes across all FAST blocks. In contrast, the initial baseline is a difficult and novel task.

The current results suggest that a number of practice blocks are required for the participant to develop a generally effective response strategy. However, if we were to introduce multiple Practice Blocks, one of the main advantages of the FAST procedure (its namesake speed of application) would be lost. In addition, providing up to seven practice blocks may have the effect of stabilising behaviour to the extent that the FAST effect is paradoxically lost due to rapid learning rates during both of the critical FAST test blocks. Indeed, this may be why no effect was observed here, even for the group of participants that had been exposed to seven practice blocks before the critical FAST blocks. Another strategy could be employed to aid participants in completing the task. The instructions used in the experiments in this chapter and the first three experiments of Chapter 2 were;

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.

While these instructions give all the necessary information to complete a FAST test block, they fail to specify some important contingencies. The consequences of a non-response, and the window in which one must respond, are not specified. Nor do the instructions specify the precise contingencies involved in "respond without error". The introduction of clearer instructions, which specify the learning rules without specifying the correct responses, would be helpful in accelerating a participants' ability to complete the procedure. In addition, the researcher can also add a "counter" which will give the participants unambiguous feedback with regard to their performance and serve as a reminder of the consequences of a non-response or incorrect response, above and beyond the Correct/Wrong feedback already provided. These improved instructions and the addition of a counter were included in Experiments 4 and 5 (which were conducted after the research described in this chapter) and in the experiments in Chapter 4.

Experiment 8: Response Criterion

To complete a FAST block, a participant must make a series of responses with minimal errors. The absence of a response within the response window is recorded as an error, and as such, both rapid and accurate responding is required in order to complete a block. In other FAST experiments, the criterion has varied either requiring ten correct responses without error, or allowing a single error in a sequence of ten trials.

These criterion were based on arbitrary decisions taken by the experimenter at various times, with the goal being that the criterion is challenging enough that it can be said that the participant is responding fluently, but easy enough that blocks can be completed quickly. The variation in criterion across experiments was part of a general inductive approach but no clear effects for such variations could be ascertained in the absence of a dedicated analysis. The initial criterion of ten correct responses in a row appeared to be effective (i.e. because expected effects were observed in experiments using this criterion). However, FAST trials-to-criterion display quite a large amount of between-participant variance. This could partially be attributed to the fact that a single error on a FAST blocks results in the participant being forced to complete a minimum of ten further trials. A participant who is fluent in the task may make 9 correct responses, followed by an error, followed by 10 more correct responses. Another participant with similar fluency may simply make 10 correct responses, and emerge with a trials-to-criterion score ten higher than his counterpart. In effect, there may be a very large difference in the trials to criterion observed using these two criteria, or it may emerge that it rarely occurs that a fluent participant produced 9 correct response followed by an error. Ascertaining this was the focus of Experiment 7. The current experiment investigates three different FAST response criteria across three different participant groups and their impact on FAST effect sizes.

Method

Participants

Thirty volunteers participated in the study as participants. Data from three participants were excluded due to exceeding 100 trials to criterion on the FAST blocks (see below). Unfortunately, due to experimenter error, demographic data for participants is not available.

Apparatus and Stimuli

All phases of the experiment were presented to participants on an Apple Macbook laptop computer with a 13-inch monitor (1024 x 768 pixel resolution). Stimulus presentations were controlled by the Psyscope software package (Cohen, MacWhinney, Flatt & Provost, 1993) which also recorded all responses. Stimuli consisted of twenty four nonsense syllables randomly assigned to their roles as samples, comparisons, and FAST test stimuli (see below). These will be referred to in future using alphanumerics.

Ethics

All participants were presented with and signed a consent form before proceeding to first phase of the experiment (See Appendix 1). Participants were told informally that performance on the task would not allow the researcher to make any individual psychological assessments but may allow for group patterns to be identified. All participants were informed of the true nature of the study after participation and were given the opportunity to withdraw their data at that stage.

General Experimental Procedure

The experiment consisted of two phases. Phase 1 (A-B Training) consisted of a matching to sample protocol to establish two simple stimulus relations (A1-B1- and A2-B2). All participants received identical stimulus-stimulus training.

Phase 2 (Function Acquisition Speed Tests) which probed for the A1-B1 relation trained in phase 1. Each FAST consisted of two test blocks (consistent and inconsistent), with four additional baseline blocks interspersed between them. (Baseline 1, Test Blocks 1, Baseline 2, Test Blocks 2, Baseline 3, Test Blocks 3). For the second phase, participants were randomly assigned to one of four experimental groups. Each experimental group were exposed to the same experimental procedure in this phase, with one difference – the criterion required to pass a FAST block (see below).

All phases were presented consecutively by the computer software, which also controlled the delivery of instructions at the beginning of each phase. Participants sat comfortably at a standard computer desk and viewed the screen at a distance of approximately 60-70cm and at eye level.

Phase 1: A-B Training

This phase was identical in all respects to Phase 1 in Experiment 6. Figure 6 describes the stimulus relations trained during this phase.

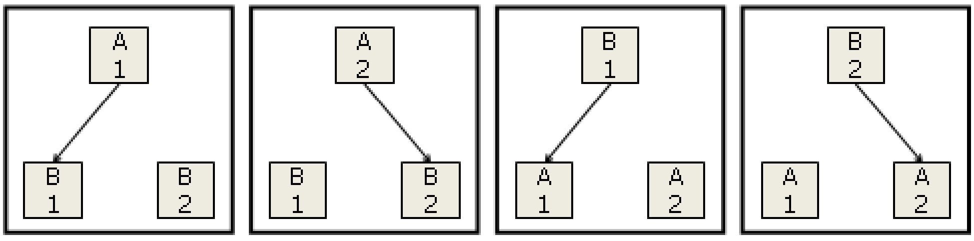


Figure 6. The stimulus relations trained during Phase 1. Solid lines indicate reinforced matching responses.

Phase 2: The Function Acquisition Speed Test

This phase was identical in all respects to Phase 3 in Experiment 6, with one exception. Participants were divided into three experimental groups. The criterion for completion of a block was varied by experimental group.

In order to complete a block, Group 1 participants were required to make no more than 3 errors in any sequence of 10 trials in order to complete a block (i.e., 7 correct out of 10). Group 2 were required to make no more than 2 errors (i.e., 8 correct out of 10), while Group 3 were required to make no more than 1 error in a sequence of 10 trials (i.e., 9 out of 10). Figure 7 below describes the block order and required responses for the FAST.

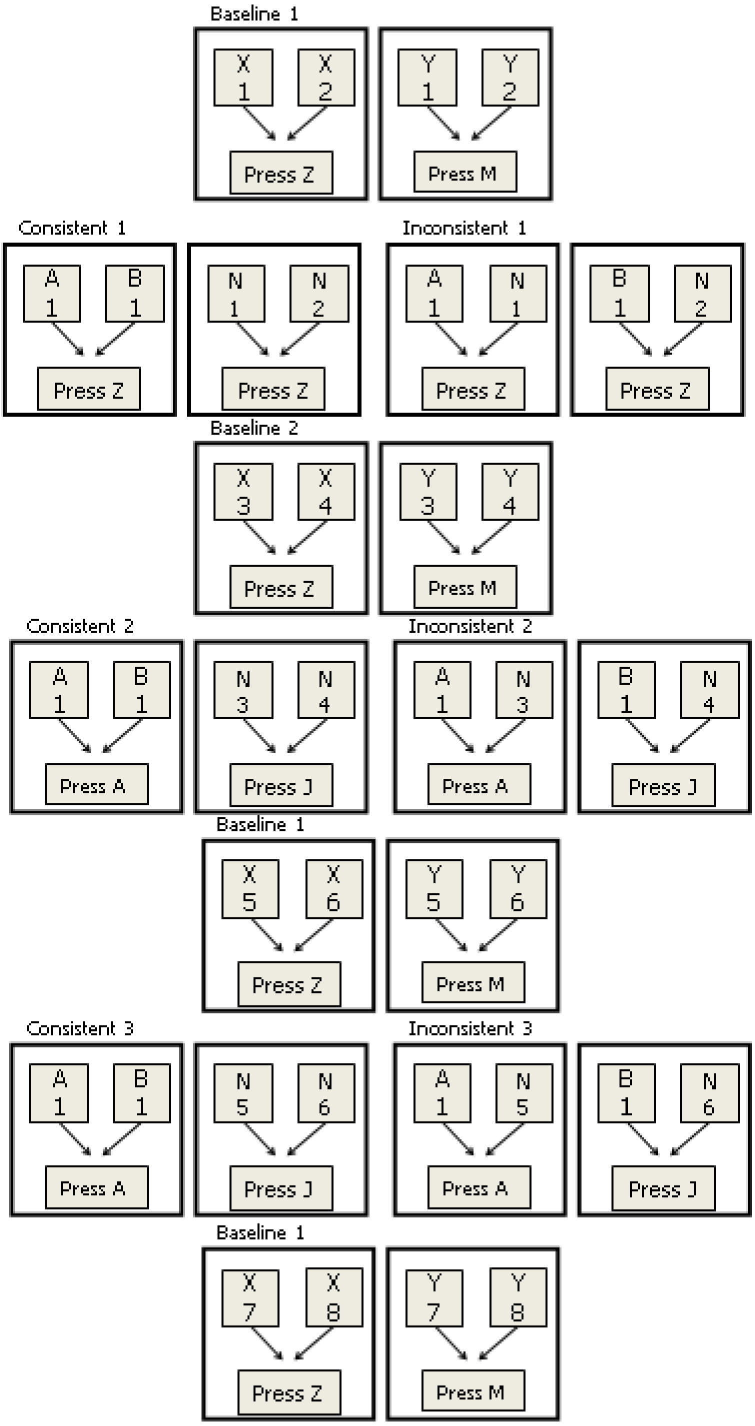


Figure 7. A schematic of the block order and various response functions assigned to stimuli during baseline and FAST blocks. The panels, read from top to bottom, represent each FAST block presented in sequence.

Results

Matching to Sample

All participants completed the same matching to sample training. For Group 1 the mean number of trials required to complete matching to sample training was 54.4 (SD = 21.598). For Group 2, the mean was 56.89 (SD = 26.667), and for Group 3, the mean was 44 (SD = 16.562).

Baseline Blocks

Participants completed four baselines over the course of the experiment. For Group 1 (7/10 Criterion), the mean baseline was 20.575 (SD = 6.035). For Group 2 (8/10 Criterion), the mean baseline was 33.472 (SD = 13.954). For Group 3 (9/10 criterion) the mean baseline was 37.468 (SD = 15.598). A one-way ANOVA showed significant differences in mean baselines between the experimental groups (F = 4.824, p = .017). Post hoc Bonferroni-corrected planned comparisons showed that Group 1 and Group 3 were significantly different from one another (p = .023). More specifically, the mean baselines for Group 1 (7/10 criterion) were significantly lower than the mean baselines for Group 3 (9/10 criterion) with a mean difference of 16.894.

Function Acquisition Speed Test

Group 1. Ten participants completed three runs of the FAST with a 7/10 block completion criterion. In the first run, only four of the ten participants showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 13.5 (SD =8.266). The mean difference in the unexpected direction was -22.333 (SD = 16.157). The overall mean difference between inconsistent and consistent blocks was -8 (SD = 22.588). The mean SOR was -2.818 (SD = 7.542). A Wilcoxon signed ranks test showed that this was not significantly different from zero (z = -.986 p = .33).

On the second run of the FAST, two out of ten participants showed a FAST effect in the expected direction. Three participants showed no difference. The mean difference between inconsistent and consistent blocks in the expected direction was 31.5 (SD = 30.406). The mean difference in the unexpected direction was -11.75 (SD = 8.958). The overall mean difference between inconsistent and consistent blocks was 1.6 (SD = 19.174). The mean SoR on the second run was 0.582 (SD = 6.344). However, this was not significantly different from zero, according to a Wilcoxon signed ranks test (z= -.105, p = .917).

On the third run of the FAST, four out of ten participants showed a FAST effect in the expected direction. One participant showed no effect. The mean difference between inconsistent and consistent blocks in the expected direction was 7.5 (SD = 1.291). The difference in the unexpected direction was -26.2 (SD = 32.568). The overall mean difference between inconsistent and consistent blocks was -10.1 (SD = 26.239). The mean SoR on the third and final run was -3.18 (SD = 9.096). A Wilcoxon signed ranks test showed no significant difference from zero (z = -1.007, p = .314).

Group 2 (8/10 criterion). In the first run of the FAST, seven of the nine participants showed a FAST effect in the expected direction. The mean difference in the expected direction was 24.429 (SD = 9.144). The mean difference in the unexpected direction was -7.5 (SD = 6.364). The overall mean difference between inconsistent and consistent blocks was 17.333 (SD = 16.309). The mean SOR was 5.084 (SD = 4.68). A Wilcoxon signed ranks test showed that this was significantly different from zero (z = 2.31 p = .021).

On the second run of the FAST, four out of nine participants showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 23 (SD = 10.165). The mean difference in the unexpected direction was -14.8 (SD = 11.122). The overall mean difference between inconsistent and consistent blocks was 2 (SD = 22.305). The mean SoR on the second run was 0.337 (SD = 6.529).

However, this was not significantly different from zero, according to a Wilcoxon signed ranks test (z = .296, p = 0.767).

On the third run of the FAST, eight out of nine participants showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 9.875 (SD =6.128). The difference in the unexpected direction was -8. The overall mean difference between inconsistent and consistent blocks was 7.889 (SD = 8.268). The mean SoR on the third and final run was 2.372 (SD = 2.613). A Wilcoxon signed ranks test showed that this difference was significantly different from zero (z = 2.073, p = .038).

Group 3 (9/10 Criterion). In the first run of the FAST, four of the eight participants showed a FAST effect in the expected direction. The mean difference in the expected direction was 16 (SD = 11.518). The mean difference in the unexpected direction was -22.75 (SD = 12.339). The overall mean difference between inconsistent and consistent blocks was -3.375 (SD = 23.476). The mean SOR was -1.141 (SD = 6.399). A Wilcoxon signed ranks test showed that this was not significantly different from zero (z = -.420, p = .674).

On the second run of the FAST, five out of eight participants showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 25.8 (SD = 22.399). The mean difference in the unexpected direction was -25.667 (SD = 6.658). The overall mean difference between inconsistent and consistent blocks was 6.5 (SD = 31.765). The mean SoR on the second run was 2.041 (SD = 9.131). However, this was not significantly different from zero, according to a Wilcoxon signed ranks test (z = .420, p = .674).

On the third run of the FAST, five out of eight participants showed a FAST effect in the expected direction. The mean difference between inconsistent and consistent blocks in the expected direction was 9.2 (SD = 7.429). The mean difference in the unexpected direction was -9.333 (SD = 4.726). The overall mean difference between inconsistent and consistent blocks was 2.25 (SD = 11.398). The mean SoR on the third and final run was .671 (SD = 3.42). A Wilcoxon signed ranks test showed that this difference was not significantly different from zero (z = .230, p = .779).

Excluded Participants

Two participants were excluded from the above data analysis on the basis of their failing to complete a FAST block before the 100 trial limit was reached. Of these two participants, 1 participant was in Group 2 (8/10), and 1 was in Group 3 (9/10).

Between Group Effects.

The SoR scores from each of the three groups on each run of the FAST were compared using a one-way ANOVA. The first run of the FAST showed a significant difference (F = 3.941, p = .033). Post-hoc Bonferroni-corrected analyses showed that SoRs for Group 1 and Group 2 were significantly different (p= .038), with a mean difference of 7.902.

The mean number of trials to criterion across all blocks for Group 1 was 20.5 (SD = 3.537), for Group 2 was 28.733 (SD = 8.003), and for Group 3 was 31.95 (SD = 9.11). Unsurprisingly, the 7/10 criterion allowed participants to complete a block more quickly than those in the other two conditions.

On the second run of the FAST, no significant differences were found (F = .133, p = .876) and the same was true of the third run (F = 2.408, p = .111).

Summary

Only Group 2, with the 8/10 criterion, showed any statistically significant FAST effects. On the first and third run of the FAST, participants in Group 2 showed a group level FAST effect in the expected direction.

Discussion

As would be expected, participants completed FAST blocks more quickly with a 7/10 criterion. SoR scores were correspondingly lower, and few participants showed differences in the direction predicted by the pre-experimental A-B training. Some of this is attributable to same reasons that applied to the previous experiments, such as insufficiently robust training and a complex order effect.

The 7/10 criterion allows participants to complete a block (and thus record a trials-to-criterion score) while allowing for 3 errors on non-responses. This is not a sufficient level of rigour to ensure that participants are fluent in the task, and completion of a FAST block may simply be too easy. In effect, a floor effect is encountered in the data and any differences in trial requirements across blocks are diminished or erased. By contrast, those participants in the 8/10 criterion group showed more significant FAST results in the predicted direction than have been observed in similar experiments in this chapter. While participants experienced a flattening of SoR scores in the second run, overall more participants performed as predicted, and with larger SoR scores.

The 9/10 criterion group showed a similar pattern of results to the 7/10 group. This is surprising, as 9/10 has been used as a criterion in FAST experiments which have been highly successful and shown the predicted patterns of results. This gives additional weight to the argument that much of the variance in the experiments in this chapter can be attributed to factors such as insufficient function training, poor experimental control, and methodological flaws such as the complex order effect.

The results of this experiment were suggestive that allowing 2 errors in a run of ten responses would produce the most reliable FAST effects. However, the “counter” innovation suggested in the previous experiment is best implemented with a “10 correct responses in a row” criterion. The benefits of the counter outweigh the potential benefits of a more lax criterion.

General Discussion

The experiments in this chapter were designed with the intention of analysing the contributions of several small procedural modifications on the outcome of a FAST. Furthermore, a series of FASTs were run in order to consider the stability of the procedure over time. To keep the task difficulty low, the pre-FAST trained relations that would be tested were simple A-B relations.

Experiment 6 varied the response window in which the fourty-four participants must respond between groups. 1s, 2s, 3s, and 4s response windows were investigated. Participants in the 4s condition showed no FAST effect whatsoever, confirming that a longer response window will undermine the process of an implicit test, as per the REC model (Barnes-Holmes et al., 2010). Participants who completed the procedure under 1s response windows produced the most consistent FAST effects – however, this group also showed the highest number of participants reaching the 100 trial limit, which resulted in the removal of their data. This result demonstrates the tension between as strict a response window as possible and the deleterious effects of a too stringent response window.

Experiment 7 introduced a number of baseline-like practice blocks. Twenty-three participants across three groups completed 3, 5, or 7 practice blocks prior to the FAST. None of the runs of the FAST produced statistically significant results on the FAST on a group level, for any of the groups. However, this experiment suffered from a low sample size, and as such, it is difficult to conclude that this was down to a deleterious practice effect. What was of interest, however, was the fact that a noticeable drop in trials to criterion required from the fourth practice block to the fifth in the two groups who completed five and seven practice blocks each.

Experiment 8 varied the criterion for block completion between the thirty participants. One group were allowed 3 errors in a run of 10 responses (7/10 criterion), the second group were permitted 2 errors (8/10 criterion) and the third were allowed only one error (9/10 criterion). The 8/10 criterion produced the most reliable FAST effects, while the 7/10 criterion produced a notable floor effect in the data, as it was insufficient to ensure task fluency.

In each of the experiments reported in this chapter, the most noticeable pattern in the results was the inconsistency in statistically significant FAST effects (i.e., positive SoRs being different from zero on a group level). Given the simple nature of the trained and tested A-B relations, and the success of Experiment 1 in detecting the same relations, this is surprising and disappointing. This effect cannot be solely attributed to practice, as in Experiment 6, all significant results emerged on the third run of the FAST, while the significant results in Experiment 8 emerged on the first and third run for Group 2. If the effects had been due merely to practice, it would be expected that performance on consistent and inconsistent blocks would converge, resulting in a loss of meaningful differences. It is possible that improved proficiency with the test format overall makes the disruptive effects of the task switching in the test blocks more disruptive than they otherwise would be, but this interpretation is undermined by all of the results in the prior chapter. As such, we must look elsewhere for an explanation.

In Experiments 1, 2, and 3, concern was expressed that a number of participants were not showing the predicted FAST effects. It was hypothesised that this was somewhat due to the fact that the trained arbitrary relations were not trained to a point where they were sufficiently robust. This may indeed be a factor in the current experiments. As noted before, Experiment 4 was conducted after the experiments in this chapter, and the results of that experiment strongly suggest this interpretation.

One major issue brought to the researcher’s attention by the experiments described herein concerns the effect of order. When a single pair of FAST test blocks was run, order had not been found to be an issue. However, when multiple blocks are run in serial, a deeper issue emerges.

At the onset of each *pair* of FAST Test Blocks, the order of that pair was randomized. This allows for a wide variety of possible order combinations across the six test blocks - 3 consistent and 3 inconsistent. For example, it is possible for the random selection to produce an order of consistent, inconsistent, inconsistent, consistent, consistent, and inconsistent. Considered as isolated pairs, this would not be a problem, but when considered in sequence the difficulties become clear. Consider, for instance, a participant who received their FAST blocks in the order described above. If they displayed an expected response in their first FAST block, this would suggest that the pre-experimental relations were well trained, and as such, the experimenter would expect this performance to maintain across multiple runs. However, if they then received 2 inconsistent blocks in a row (separated by Baseline 2, in which the contingencies are not relevant to the test task), rapid responding in that block would already have been established during the previous exposure to an inconsistent block with those same stimuli. When given yet another inconsistent block shortly thereafter, fluency would be established quickly. Following this, when the subsequent consistent block begins, the prevailing contingencies would now inconsistent with those that have been in effect for the previous two blocks. In a strange reversal of what was intended, the consistent block may at least momentarily become more difficult than the inconsistent block. The same reversal could occur in the transition from FAST 2 to FAST 3. In addition, this possible process may have had particularly strong effects in the current research compared to that using real word stimuli insofar as participants had received limited training, which may not have been robust enough to withstand the demands of regular contingency switching in the manner described above.

The complex issue of block order when multiple FASTs are delivered is only compounded by the fact that there are eight potential combinations of orders produced by three consecutive tests. In a study with a larger sample size, it may be possible to analyze the effects of each distinct sequence more carefully, but in a study of this size in which participants from different groups cannot be compared, such an analysis is impossible.

In future experiments involving multiple pairs of FAST blocks, if counterbalanced order is necessary, the entire sequence of test blocks should be in the order “consistent first” or “inconsistent first”. This preserves the task switching dynamic and avoids any two consecutive blocks with the same response contingencies.

Several of the changes suggested by the foregoing experiments were incorporated into the FASTs which followed. A single practice block was included in future FASTs, allowing some degree of practice but avoiding the deleterious potential effects of multiple practice blocks. On each trial in future FASTs, there was a response window of 2000ms. The criterion for completion of a block of trials was ten correct responses in a row. This one change runs counter to the criterion suggested by Experiment 8, and this decision was taken due to the inclusion of a “counter” which displayed the number of correct responses the participant had emitted without error. Further, the instructions presented to the participants at the onset of each FAST were clarified, and participants were required to wait a minimum of 30 seconds before proceeding from the instructions page to the test itself. It was judged that these modifications retained the important balance of difficulty necessary for an implicit test to operate at maximum efficiency.

Further research into the precise configuration of the FAST procedure will of course be valuable and necessary to the continued development of the tool. However, as this is the first piece of work on the topic, it was considered more prudent to take the tool outside the laboratory for some final experiments test its functionality with regard to naturally occurring stimulus relations. Ultimately, the FAST is a tool intended to be used in the real world, and as such it was necessary to test a pilot version of the tool in an applied context and to put aside for now the idea of “perfecting” the method in the context of a single research programme.

Chapter 4

Using the FAST to assess real world histories of verbal behaviour

Traditionally, "implicit" test methodologies have been developed with research in which only real words taken from the vernacular are used as stimuli. However, by beginning with stimulus relations entirely under experimental control in the current research, it has been possible to avoid potential confounds relating to real-word stimulus choice, word frequency, word length, and so on. Thus, the basic FAST methodology is a fruitful avenue of research for further development, and the FAST can be used with some degree of confidence in future research. Most importantly for now, however, the FAST procedure was shown to be sensitive to laboratory trained relations and derived relations that had never been reinforced in the history of the participants.

In Chapter 2, the FAST methodology was developed and tested under strict laboratory conditions. The stimuli were arbitrary nonsense syllables, and the participants' histories of relating those stimuli were entirely under the experimenter's control. In developing a methodology designed to detect a phenomenon, the most prudent course of action is to test that methodology in conditions where that phenomenon is known to be likely to occur. An appropriate analogy would be testing a new type of radiation detecting device. In initial testing, the developer would expose the device to a source known to be radioactive, to ensure that it was capable of detecting radiation in a situation where its occurrence is highly probable. The FAST methodology has passed this test, proving to be capable of detecting known histories of relating arbitrary stimuli, particularly when these histories are well established.

The next challenge for the FAST methodology is to test it in "real world" situations, where the histories it is attempting to detect are not under the control of the experimenter. That is to say, we must attempt to put the FAST to use in detecting relations between real words in the histories of the participants. Other implicit testing methodologies used real words as stimuli in their very first experiments (IAT; Greenwald et al., 1998, IRAP; Barnes – Holmes et al., 2006). However, this has left the underlying processes of the tests in doubt (particularly in the case of the IAT), and effects remain based on untested assumptions that reaction time is a valid measure of the hypothesised construct (i.e., attitude). Any given participant has a complex history with the use of any particular word, particularly those that relate to socially sensitive or powerfully emotive topics – the very topics that implicit testing methodologies are designed to deal with. As such, launching directly into analyses using real word stimuli without first confirming the "proof of concept" for the FAST would have been premature. Now that the basic processes have been validated, it is possible to begin testing the FAST in the context of real world "attitudes".

As noted in Chapter 1, attitudes can be conceived of an instance of verbal behaviour (or a series of related verbal behaviours) that emerge from a history of relating relevant verbal stimuli according to which the functions of events are transformed. From this perspective, the IAT (and the IRAP) are measures of a participants' verbal fluency with regard to different pairs of stimuli in a task switching context. The particular form of the IAT and IRAP tasks also means that they detect subtly different verbal behaviours – and the same is true of the FAST format. This is the case even though some fundamental processes may be shared in the performance of all three test formats.

One noted feature of the IAT is that it provides a relative measure of association strength (De Houwer, 2002, Greenwald, Nosek & Banaji, 2003). Each trial of the IAT asks the participant to categorize a stimulus as either (for example) "black/bad" or "white/good". When attitudinally coherent items share a response key, a participant is assumed to respond more quickly than when the categories are counter-attitudinal (i.e. "black/good" and "white/bad"). However, due to the fact that an IAT trial involves both responding towards one category while responding away from the other, it is impossible to determine which of these cues is at work in controlling the participants' response, or the extent to which one or the other cue contributes to the total effect.

In a "race" IAT, a positive D-score indicates a pro-white bias. This result indicates that participants respond more quickly on trials where white/good and black/bad each share a response and more slowly on trials where black/good and white/bad each share a response. This result, however, could reflect rapid responding to white/good more so than black/bad (a pro-white bias) or more rapid responding to black/bad trials (an anti-black bias). The IAT methodology does not allow the researcher to determine which of these is the case. Thus, the IAT offers information only on the question; "Is White better than Black" without specifying whether or not the test outcome derives from a participant having a history of pro-white responding or a history of anti-black responding, or indeed a lack of a history of pro-black responding – all of which could indicate quite different "attitudes" and expected behavioural outcomes. This has been cited as a severe limitation of the IAT, as it creates ambiguity with regard to the precise history being detected (or, in social cognitive terms, the form of the attitude being detected). As such, the IAT provides only indirect evidence for a particular set of beliefs, by providing an index of associations that are probably involved in that particular attitude (De Houwer, 2002)

The IRAP seeks to address this limitation by measuring beliefs more directly (Barnes-Holmes et al., 2010). The IAT does not offer the participant any opportunity to respond to the specific verbal relations thought to underlie their beliefs. By contrast, each trial of the IRAP presents a specific relation between a label stimulus and a target item, which the participant is required to confirm or deny based on the response criteria of the particular block. Participants are asked to either confirm or deny a proposition that is counter attitudinal in the inconsistent blocks, and to either confirm or deny a proposition that is pro-attitudinal during the consistent blocks. This allows the IRAP to index the associations between individual stimulus items. For example, in a race IRAP, a D score is produced for black-good, white-good, black-bad, and black-good attitude dimensions The IRAP is therefore a non-relative measure of relations between stimuli, capable of isolating which particular pair (or pairs) of relations is producing the overall D score.

While the IRAP allows the researcher to separate out individual pairs of relations between the label stimuli ("good" and "bad") and the two sets of target stimuli (black face and white faces), it's structure still requires that the attitude in question be somewhat dyadic. In order to see an "IRAP effect", one must be able to compare response latencies on inconsistent trials with response latencies on consistent trials. This requires a symmetrical arrangement involving two labels and two targets such as "black/white" and "good/bad".

Two IAT variants also attempt to address this issue; the Single Target IAT ST-IAT; Wigboldus, 2004) and the Go/No Go Association Test (GNAT; Nosek & Banaji, 2001). In the single target IAT, only one target category is used, and participants are asked to respond to this target (e.g. "black") and one attribute (e.g. "bad") with one response key, and to the other attribute (i.e. "good"). In the GNAT, participants are required to "go" (i.e. respond with a key press) when presented with stimuli that belong to the target category or attribute. (e.g. "white" and "good") and to make no response (no go) when the stimulus belongs to neither. In both of these cases, the attribute label must again be dyadic. While it improves upon the issues facing the IAT (which must be dyadic on two axes), these efforts do not circumvent the issue entirely. The FAST offers a researcher the ability to test for a one single relation, quickly and simply, without the requirement to simultaneously test for the inverse attitude.

This is useful in the sense that in the context of creating behaviour change, we are not often concerned with the counter-attitude. For example, when trying to treat spider phobia, it is not necessary to cause the participant to respond to arachnids as appetitive, but simply to reduce the aversive functions of the stimulus We need only know that the relation "spiders-bad" is operating, and that after treatment, the relational no longer appear to be functioning. A positive spider bias is not necessary for successful treatment, and a measure which is contaminated with those elements of a relational network may be misleading.

The following chapter describes three experiments in which the FAST is used with real world stimuli in order to detect particular verbal histories across groups of participants. In Experiment 9, the participants are exposed to a FAST procedure using real word stimuli, in which the test stimuli are “spider” and “disgust”. Participants are also asked to complete an “explicit” pen and paper measure to assess their public attitude to spiders, and the relationship between the questionnaire measure and the results of the FAST procedure are explored. In Experiment 10, a similar experiment is conducted in which the test stimuli are the words “immigrant” and “cheat”. Again, the results of the FAST are compared to the results of a questionnaire measure. In Experiment 11, a series of FAST procedures are employed in a repeated measures design. Each FAST measures the relation between sexual images and images of young females in a heterosexual male population. In each of the three FASTs, the age of the young females is varied. This allows for an examination of the age at which female images and sexual images become easily related.

Experiment 9

Measuring naturalistic stimulus relations 1: Spiders and Disgust

Spider phobia has been historically linked to a predator defence mechanism, wherein phobic individuals fear the physical harm that can be caused by small venomous animals (Ohman, Dinberg, and Ost, 1985). In more recent times, it has been suggested that disgust may be an important factor in spider phobia – in the sense that disease avoidance is the primary motivator of fearful avoidance reactions (de Jong, Peters & Vanderhallen, 2002). Spider fear often co varies with avoidance of animals which are considering disgusting rather than potentially physical harmful, such as maggots (Davey, 1992). In one study by de Jong et al. (2002), sensitivity to contagion was the best predictors of phobic group membership.

Researchers then turned to the use of implicit measures to clarify the contribution of disgust to spider fear further. Spider fear is an example of what it termed "non-cognitive" fears, in the sense that spider phobic individuals often have difficultly verbalizing the source of their fear, and spiders are often spoken about as intrinsically fearful by such individuals (de Jong, van den Hout, Rietbroek, Huijding, 2003). This makes spider fear an excellent candidate for investigation via indirect methods, such as the IAT and the FAST.

Teachman and Woody (2001) were able to distinguish between phobic and non phobic groups using an IAT paradigm. Using both the IAT and the affective Simon paradigm, de Jong and colleagues (2003) tested for negative associations with spiders in high and low fear individuals, and found that both groups showed significant negative automatic associations. By contrast, de Jong and Huijding (2007) found that both a threat IAT and a disgust IAT distinguished between phobic and non phobic individuals. The equivocal results in these experiments could potentially be attributed to exemplar choice. More specifically, the IAT requires the use of dyadic stimulus pairs. In the Teachman and Woody (2001) study, spiders were contrasted with "snakes". De Jong et al. (2003) contrasted spiders with a neutral concept, and De Jong and Huijding (2007) contrasted spiders with weapons (in the "threat" IAT) and with maggots (in the disgust IAT). The different results found in these studies demonstrate the difficulty of analysing some concepts using the IAT framework, as a result of the need for a contrast category. What is clearly needed in such research contexts is a single-target test such as the FAST. The FAST paradigm is designed to allow for the analysis of a single relation between stimuli, which, given the difficulties articulated in the above cited papers, would be of interest to researchers in this field.

In the following experiment, participants completed a FAST to assess the strength of relation between the test stimuli "spiders" and "disgust" and its correlation to the "Fear of Spiders Questionnaire" pen and paper measure. All neutral stimuli were also real words. In line with predictions made in other implicit studies of spiders disgust, it was expected that high SoR scores for the spiders/disgust FAST would successfully identify participants who reported high fear of spiders and also that a generally high level of spider-disgust associations should found in the general population,

Method

Participants

Eighty-eight participants participated in the current study. Ages ranged from 18 to 25 (Mean 20.45, SD = 1.371). Forty-eight participants were male, forty were female. Participants were recruited from acquaintances of the experimenter and from the undergraduate population at NUI Maynooth.

Apparatus and Stimuli

The FAST blocks were presented to participants on an *Apple Macbook®* laptop computer with a 13-inch monitor (1024 X 768 pixel resolution). Stimulus presentations were controlled by the software package *Psyscope* (Cohen, MacWhinney, Flatt & Provost, 1993) that also recorded all responses. Stimuli consisted of fourteen words;"rug", "tree", "board", "lamp", "coin", "fork", "paper", "van", "spider", "disgust", "plaster", "margin", "crate", "door", "hand" and "ground."

Participants were also presented with the Fear of Spiders Questionnaire (FSQ) created by Szymanski and O'Donohue (1995) to complete following the FAST procedure. This consists of 18 fearful statements regarding spiders with each item rated on an eight point scale ranging from 0 (totally disagree) to 7 (totally agree). The FSQ has a high test retest reliability (α = .97). The FSQ is capable of differentiating between individuals with and without spider phobias and is sensitive to the effects of treatment (Szymanski & O'Donohue, 1995). It is also correlated with scores on the Spider Questionnaire as well as ratings during a behavioural avoidance test (Szymanski & O'Donohue, 1995). At the same time that participants completed the FSQ, they also completed the *Marlowe-Crowne Social Desirability Scale.* (M-C SDS; Crowne & Marlowe, 1960) The SDS consists of 33 items which measure participants' need to respond in "culturally sanctioned ways". Participants must respond true or false to each of these statements. The SDS correlates significantly with the “Lie” scale and the “K” (test taking attitudes) scale on the MMPI. The SDC has high test retest reliability ($\alpha = .89$).

Ethics

All participants were presented with and signed a consent form before proceeding to first phase of the experiment (See Appendix 1). Participants were told informally that performance on the task would not allow the researcher to make any individual psychological assessments but may allow for group patterns to be identified. All participants were informed of the true nature of the study after participation and were given the opportunity to withdraw their data at that stage.

General Experimental Procedure

In this experiment, participants were exposed to a FAST procedure designed to probe for the existence of a history of relating “spiders” with “disgust”. The FAST consisted a of a practice block, two test blocks and two baseline blocks, with one administered before the test blocks, and one afterwards. Unlike previous experiments, all stimuli used in the FAST procedure were real words. Figure 1 below describes the stimuli and reinforced responses for each block of the FAST, as well as displaying the block presentation order.

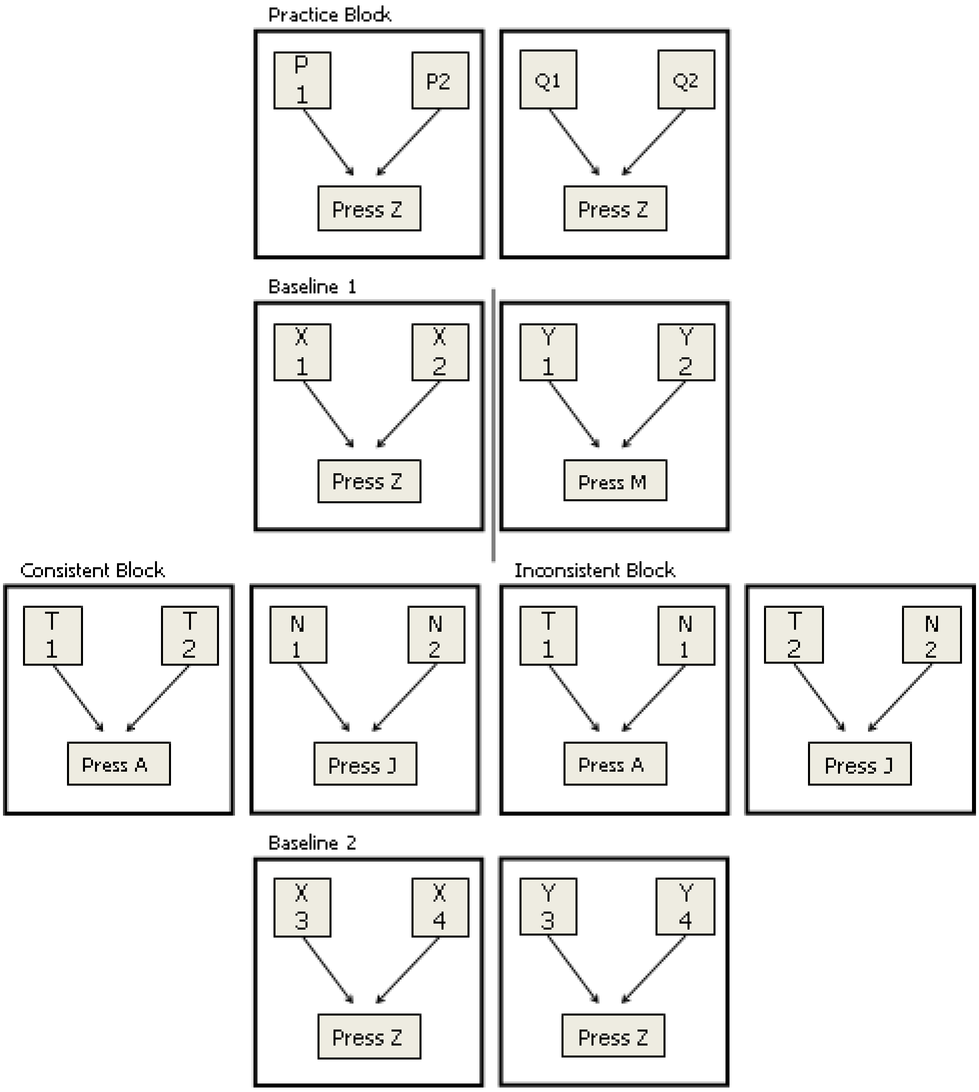


Figure 1. A schematic of the various response functions assigned to stimuli during practice, baseline and FAST blocks. The four rows, read from top to bottom, also represent the sequence of presentation of the practice block, the two baselines blocks and the two FAST blocks.

Practice Block

The practice block utilized four real word stimuli – "tree", "rug" board, and "lamp". (P1, P2, Q1, Q2). These stimuli were selected to be emotionally neutral and to be largely unrelated. The practice block was identical in format to the Baseline Blocks described below. The practice block was presented before the FAST and the baseline blocks. Participants were presented with the following instructions page at the onset of the Practice Block.

THE FOLLOWING SECTION IS A **PRACTICE BLOCK**

IN THE FOLLOWING SECTION

A single word will appear on screen

For Example

“fork”

Each time you see the word you must press either

The “**Z**” key or the “**M**” key

After you press a key, the computer will tell you if you were

CORRECT or WRONG

If you don’t press anything within 2 seconds of seeing a word, the computer will count this as a WRONG answer and move on to the next word

Your task is to learn which key **(Z or M)** to press for each of the words. This section will continue until you can make 10 CORRECT RESPONSES in a row WITHOUT ANY MISTAKES

Instructions remained fixed on screen for a period of 10s, after which time words “When you are ready to start the next section, Press Any Key” appeared at the bottom of the screen in 30pt font. When the participant pressed a key, the block began.

Baseline Blocks:

Two baseline blocks were presented (one before the first FAST testing, and one after the final FAST block). The Baseline blocks involved real word stimuli. Baseline 1 employed “fork”, “coin”, “van” and “paper” (X1, X2, Y1, Y2) as stimuli, while Baseline 2 employed “door”, “crate”, “hand”, and “ground” (X3, X4, Y3, Y4) as stimuli. These stimuli were selected to be emotionally neutral and to be largely unrelated. These phases attempted to establish the following relations, respectively; X1-Y1, X2-Y2 and X3-Y3, X4-Y4.

After the first Baseline Block was completed by the participant, a screen appeared with the following text:

Well done! You have completed this section of the experiment.

Please take a short break to rest your eyes from the computer screen.

The instructions for the next section will be displayed shortly.

This text remained on screen for 30s, after which the following instructions were displayed:

IN THE FOLLOWING SECTION

A single word will appear on screen

For Example

“fork”

Each time you see the word you must press either

The “**Z**” key or the “**M**” key

After you press a key, the computer will tell you if you were

CORRECT or WRONG

If you don’t press anything within 2 seconds of seeing a word, the computer will count this as a WRONG answer and move on to the next word

Your task is to learn which key **(Z or M)** to press for each of the words. This section will continue until you can make 10 CORRECT RESPONSES in a row WITHOUT ANY MISTAKES

Instructions remained fixed on screen for a period of 10s, after which time words “When you are ready to start the next section, Press Any Key” appeared at the bottom of the screen in 30pt font. When the participant pressed a key, the block began.

Test Blocks

The test blocks utilized the two test stimuli – “spider” and “disgust” (T1 and T2), and two further unrelated, neutral real words – “plaster” and “margin” (N1 and N2). The first of these blocks (consistent) established two functional response classes (T1-T2 and N1-N2) that would be consistent with relations predicted given a history of spider-phobic responding on the part of the participant. The second test block established two functional response classes that were inconsistent with those relations (T1-N1 and T1-N2). Unlike previous experiments, it is impossible to determine *a priori* which of these two sets of relations are actually “consistent” with the participants’ history – the naming of these blocks as “consistent” or “inconsistent” is merely an experimental convenience. As such, the order of block presentation was randomised across participants.

After the first FAST Block was completed by the participant, a screen appeared with the following text:

Well done! You have completed this section of the experiment.

Please take a short break to rest your eyes from the computer screen.

The instructions for the next section will be displayed shortly.

This text remained on screen for 30s, after which the same instructions as above were displayed. Instructions remained fixed on screen for a period of 10s, after which time words "When you are ready to start the next section, Press Any Key" appeared at the bottom of the screen in 30pt font. When the participant pressed a key, the other second test block began.

The presentation of stimuli and feedback during all blocks was identical to Experiment 4.

Participants were exposed to trials until a criterion of 10 correct in a row was reached. An upper limit of 200 trials was set.

Pen and Paper measures.

Upon completion of all of the FAST blocks, participants were then asked to complete the Fear of Spiders Questionnaire. Participants also completed the Marlowe-Crown Social Desirability Scale at this time.

Results

Function Acquisition Speed Test

Participants completed a practice block, two baseline blocks (one before the test blocks, and one after the test blocks had been completed). Table 1 below shows the mean trials to criterion for each block completed by the participants.

There was a marked decrease in the mean trials to criterion between Baseline 1 and Baseline 2, suggesting a practice effect. A paired samples t-test showed that this difference was statistically significant ($t = 3.844$, $p < 0.001$).

A further paired samples t-test demonstrated that the difference between trials to criterion on Consistent blocks (Mean = 32.43, SD = 27.63) and Inconsistent Blocks (Mean = 40.77, SD = 36.484) was also statistically significant ($t = 2.533$, $p = .013$), indicating a history of relating spiders with disgust on a group level.

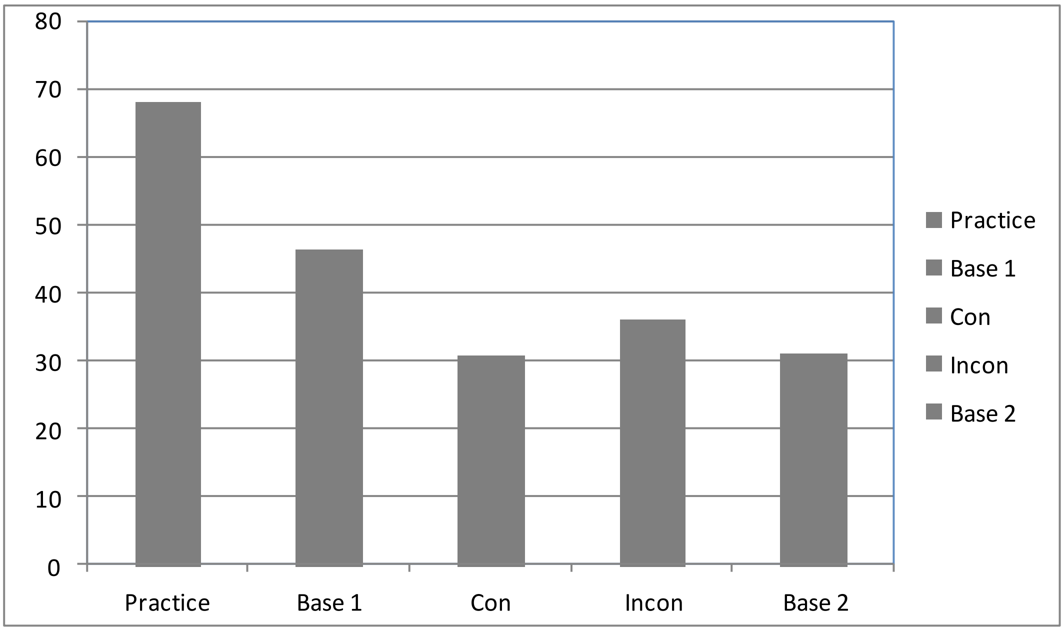


Figure 2: Mean trials to criterion for each block of the FAST. Note: the order of the consistent and inconsistent blocks was randomised across participants.

As in previous experiments, an SoR index was calculated. The mean SoR index was 2.493 (SD = 8.7534). A one sample t-test showed that SoR was significantly different from zero (t=2.672, p = 0.009) for the entire group. Due to the significant practice effect from Baseline 1 to Baseline 2, the scores were also tested for an order effect with an independent samples t-test. The test showed no effect based on order (t = -1.11, p =0.27).

Participants also completed a Fear of Spiders questionnaire. The mean FSQ score was 24.97 (SD = 27.666). The results were highly negatively skewed, indicating a low general fear of spiders. The data were then sorted into “High Fear” and “Low Fear” based on the participants’ scores on the FSQ pen and paper questionnaire. The data were split along the median (Median Fear of Spiders = 10). Table 2 below shows the mean SoR for each group. “High Fear” participants had a mean SoR of 2.939 (SD = 9.168) while “Low Fear” participants had a mean SoR of 2.025 (SD = 8.38). An independent samples t-test showed this difference to be non significant (t = .487, p =.627).

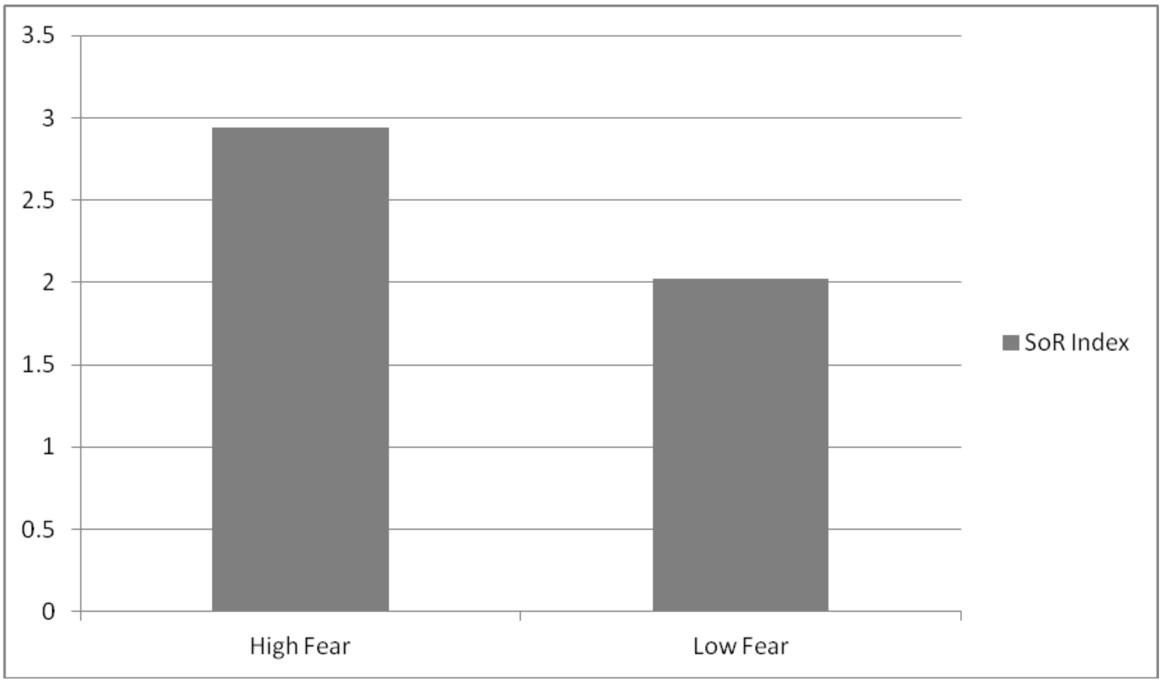


Figure 3: SoR indices for High Spider Fear and Low Spider Fear Groups

It is possible that any effects due to high/low fear of spiders were mediated by social desirability. Therefore, participants were again split by their scores on social desirability. The mean social desirability was 14.19 (5.872) and was normally distributed. The groups were split on the median (Median Social Desirability = 14). A univariate ANOVA was conducted to search for main and interaction effects for social desirability and fear of spiders. Figure 4 below graphs the estimated marginal means for each group. No significant effects for either variable on SoR were detected.

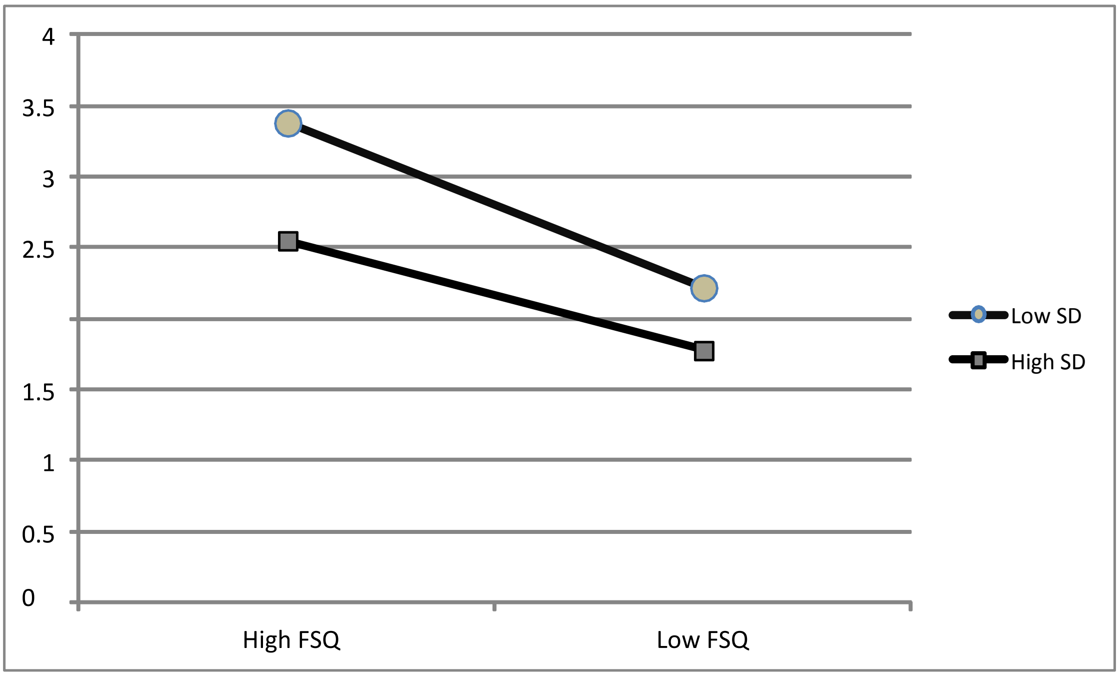


*Figure 4: Mean SoR for High/Low Fear of Spiders for High Social Desirability and Low Social Desirability Groups***.**

In summary, the FAST was sensitive to a *spider-disgust* relation in the sampled participants. The explicit measures (FSQ questionnaire) showed a low overall level of spider fear in the sample. Social desirability did not impact upon the observed results.

Discussion

The FAST test for strength of relation between “spiders” and “disgust” indicated that the two stimuli had a history of being related on a group level. However, the group level of fear of spiders was, overall, skewed towards low scores. This indicates that while spiders are widely spoken about as disgusting, that particular verbal history does not entail a corresponding fear of spiders in those that do so. It was expected that participants who showed a high SoR index for spiders/disgust would show a correspondingly high score on the FSQ questionnaires, but this prediction did not match the data. This indicates that high spider fear individuals related “spiders” and “disgust” no more than low spider fear individuals did. As a result, the FAST was unable to distinguish between high and low spider fear groups on the basis of a relation between “spiders” and “disgust”.

The results indicated an overall low fear of spiders in participants. This may be as a result of a self selection bias which removed people with a strong fear of spiders from the participant pool pre-experimentally. Few participants in the sample responded to the questionnaire in such a way as to indicate a strong fear of spiders, and as such, there may not have been sufficient variation in the sample for the FAST to distinguish between groups. However, the FAST was sensitive to a group level effect that was statistically significant, and as such the effect requires explanation.

These results do bear some similarity to deJong et al. (2002), in which an IAT showed similar negative associations for both phobic and non-phobic individuals when a neutral comparison was used. De Jong argues that in the absence of strong individual associations, participants turn to cultural stereotypes as being the only available representations. This coheres with the behavioural model of the IAT (Gavin et al., 2008) which models the IAT as a measure of relative verbal fluency. That is to say that the socio-verbal community is the source of control over the participants’ responses, insofar as they are more fluent with the vernacular, which includes the statement “spiders are disgusting”.

The use of words as stimuli may provide a reason for the way in which current results diverged from the predicted outcome. The behavioural model of the IAT describes the IAT as a measure of participants' relative verbal fluency with the tested categories. The FAST operates on similar principles, and could be thought of as measuring verbal fluency. Speaking of "spiders" as "disgusting" is common in the vernacular, even amongst non-phobic individuals. However, spider phobia involves rapid, non-verbal responses to spider stimuli, which is why the current FAST utilized verbal stimuli. It is possible that many of the functional relations of interest in this context are not common among words. In other words, it may require other forms of stimuli, such as images, to more effectively tap into the types of relations of interest here. Furthermore, it may be necessary to explore more commonly held emotional responses than spider fear, if a sufficiently large sample of participants is to be identified in a non-clinical population. Finally, it may be necessary to examine stimulus relations in a domain that is even more socially sensitive than spider fear. These measures should allow for a great probability of interesting divergences emerging between self reports and FAST outcomes, as well as providing a more reliable basis for FAST effects in the first instance. With this possibility in mind Experiment 10 will examined the more socially sensitive issue of racism and Experiment 11 employed the FAST procedure to examine relations associated with sexual preference using pictorial stimuli.

It must be kept in mind that the FAST successfully identified a history of relating spiders to disgust on the group level. This should be considered a success for this new behaviour-analytic test format. The FSQ results demonstrated that there were few individuals in the sample who displayed a high level of spider fear, and as such, it should be expected that the FAST would fail to distinguish between the contrived "high" and "low" spider fear groups. Nevertheless, a greater degree of control is still required over test outcomes and utility and this endeavour progresses with Experiment 10.

Experiment 10

Measuring naturalistic stimulus relations part 2: Ethnic Prejudice

Experiment 9 used the FAST to assess the relation between spiders and disgust in an undergraduate population. Experiment 10 ran parallel to the previously reported experiment, and adopted a similar experimental design. The current experiment seeks to use the FAST to assess socially sensitive relations regarding ethnic stereotyping, with the test stimuli being the words “immigrant” and “cheat”.

One previous study has used implicit measures to assess racial attitudes amongst Irish college students. Barnes-Holmes, Murphy, and Barnes-Holmes (2010) showed pro-black stereotyping in results from an IRAP with a 3s response window, and anti-black stereotyping when the response window was reduced to 2s. The authors interpreted this result as demonstrating that shorter response latencies more accurately tapped into underlying racist attitudes in Ireland. The IRAPs in that study used the categories black/white and safe/dangerous. However, in the years 2005-2010, the majority of immigrants into Ireland came from countries which joined the EU between the years 2004 and 2007, and negative ethnic stereotyping in Ireland is suffered by white immigrants as much as those of other ethnicities. The current study does not look at racial stereotyping in terms of skin colour, but instead focuses on immigrants in general, rather than a specific racial group. That immigrants are cheats (in both the sense of welfare fraud and in their dealings with banks) is a common negative stereotype in Ireland so those two words we selected as test stimuli. General negative adjectives such as “bad” were deemed to non-specific, and the safe/dangerous dichotomy was thought to be more specific to negative stereotyping of dark skinned people, particularly African Americans, while the stereotyping of immigrants in Ireland seems to be based more on economic conditions and conflicts over jobs and welfare.

Method

Participants

Eighty two participants participated in the study. Of these, 54 participants were male and 28 participants were female. Ages ranged from 18 to 29 (Mean 21.21, SD = 1.505) Participants were recruited from acquaintances of the experimenter and from the undergraduate population at NUI Maynooth.

Apparatus and Stimuli

The FAST blocks were presented to participants on an *Apple Macbook*® laptop computer with a 13-inch monitor (1024 X 768 pixel resolution). Stimulus presentations were controlled by the software package *Psyscope* (Cohen, MacWhinney, Flatt & Provost, 1993) that also recorded all responses. Stimuli consisted of fourteen words;"rug", "tree", "board", "lamp", "coin", "fork", "paper", "van", "spider", "disgust", "plaster", "margin", "crate", "door", "hand" and "ground."

Participants were provided with the Discrimination (DV) and Diversity scales (DS) created by F which were delivered together. Participants were instructed to indicate on a 5-point scale their agreement or disagreement on a total of 14 statements, with the ratings from 1 (strongly agree) to 5 (strongly disagree). The DS scale consisted of 10 statements regarding opinions and beliefs about discrimination in Irish society. The DV scale consisted of four statements and focused on beliefs about the value of ethnic diversity in society. The questionnaires were scored such that 1 or 2 indicated negative racial stereotyping, 4 or 5 indicated positive racial stereotyping, and 3 indicated no stereotyping. This experiment also used the Marlowe-Crowne SDS (Marlowe-Crowne, 1960)

Ethics

All participants were presented with and signed a consent form before proceeding to first phase of the experiment (See Appendix 1). Participants were told informally that performance on the task would not allow the researcher to make any individual psychological assessments but may allow for group patterns to be identified. All participants were informed of the true nature of the study after participation and were given the opportunity to withdraw their data at that stage.

General Experimental Procedure

In this experiment, participants were exposed to a FAST procedure designed to probe for the existence of a history of relating “immigrants” with “cheat”. The FAST consisted a of a practice block, two test blocks and two baseline blocks, with one administered before the test blocks, and one afterwards. Unlike previous experiments, all stimuli used in the FAST procedure were real words. After the presentation of the FAST, participants were then asked to complete two pen and paper measures - a social desirability scale and a discrimination scale. Figure 4 below shows the stimuli and response functions for each block of the FAST procedure, as well as detailing the block order.

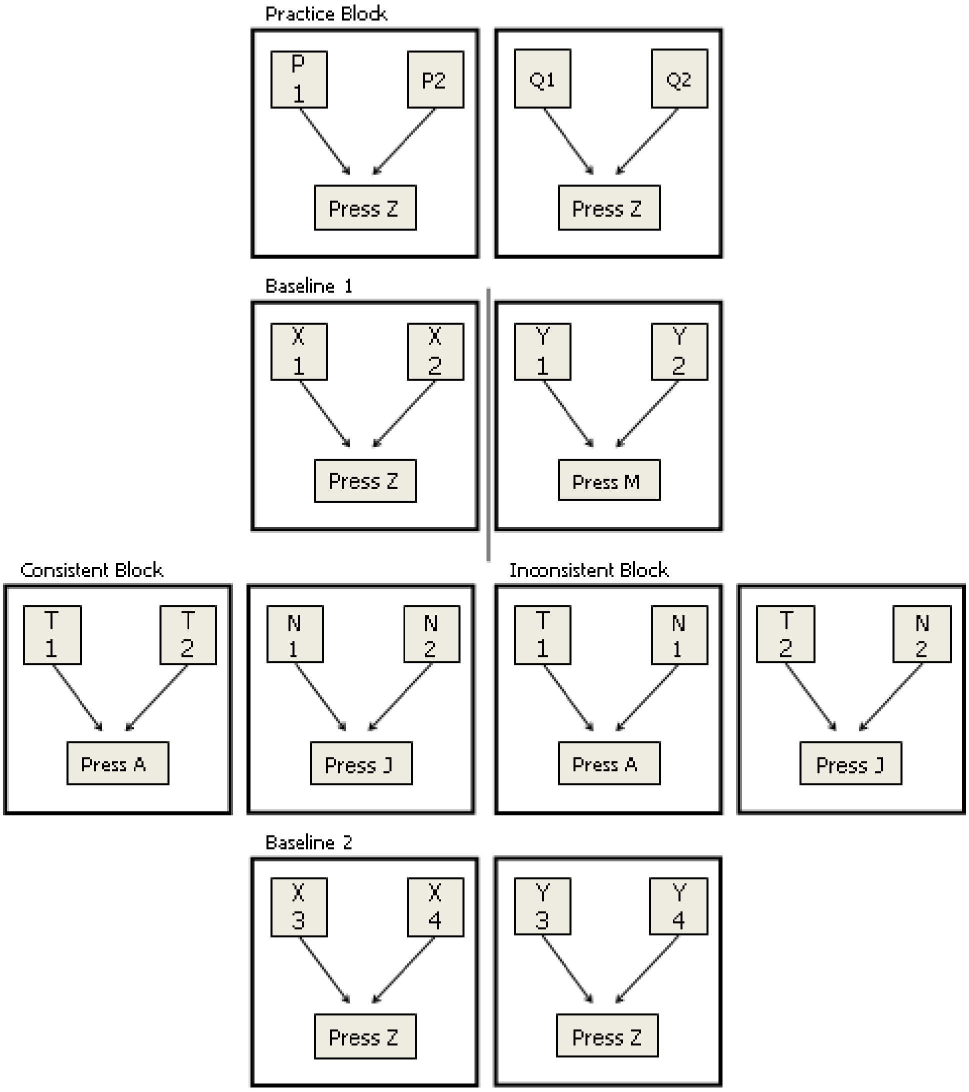


Figure 4. *A schematic of the various response functions assigned to stimuli during practice, baseline and FAST blocks. The four rows, read from top to bottom, also represent the sequence of presentation of the practice block, the two baselines blocks and the two FAST blocks.*

Practice Block

The practice block utilized four real world stimuli – “tree”, “rug” board, and “lamp”. (P1, P2, Q1, Q2). These stimuli were selected to be emotionally neutral and to be largely unrelated. The practice block was identical in format to the Baseline Blocks described below. The practice block was presented before the FAST and the baseline blocks.

The instructions presented at the start of this block were identical to those used in the practice block for Experiment 9. Instructions remained fixed on screen for a period of 10s, after which time words “When you are ready to start the next section, Press Any Key” appeared at the bottom of the screen in 30pt font. When the participant pressed a key, the block began.

Baseline Blocks

Two baseline blocks were presented (one before the first FAST testing, and one after the final FAST block). The Baseline blocks involved real word stimuli. Baseline 1 employed “fork”, “coin”, “van” and “paper” (X1, X2, Y1, Y2) as stimuli, while Baseline 2 employed “door”, “crate”, “hand”, and “ground” (X3, X4, Y3, Y4) as stimuli. These stimuli were selected to be emotionally neutral and to be largely unrelated. These phases attempted to establish the following relations, respectively; X1-Y1, X2-Y2 and X3-Y3, X4-Y4.

After the first Baseline Block was completed by the participant, a screen appeared with the following text:

Well done! You have completed this section of the experiment.

Please take a short break to rest your eyes from the computer screen.

The instructions for the next section will be displayed shortly.

The instructions presented at the start of this block were identical to those used in the baseline and test blocks for Experiment 9. Instructions remained fixed on screen for a period of 10s, after which time words “When you are ready to start the next section, Press Any Key” appeared at the bottom of the screen in 30pt font. When the participant pressed a key, the block began.

Test Blocks

The test blocks utilized the two test stimuli – “immigrant” and “cheat” (T1 and T2), and two further unrelated, neutral real words – “baker” and “pedestrian” (N1 and N2). The neutral stimuli were chosen such that all stimuli were adjectives which could be applied to any person. The first of these blocks (consistent) established two functional response classes (T1-T2 and N1-N2) that would be consistent with relations predicted given a history of discriminatory responding on the part of the participant. The second test block established two functional response classes that were inconsistent with those relations (T1-N1 and T1-N2). As in Experiment 9, it is impossible to determine *a priori* which of these two sets of relations are actually “consistent” with the participants’ history – the naming of these blocks as “consistent” or “inconsistent” is merely an experimental convenience. As such, the order of block presentation was randomised across participants.

After each FAST Block was completed by the participant, a screen appeared with the following text:

Well done! You have completed this section of the experiment.

Please take a short break to rest your eyes from the computer screen.

The instructions for the next section will be displayed shortly.

This text remained on screen for 30s, after which the following instructions were displayed:

The instructions presented at the start of these blocks were identical to those used in the baseline and test blocks for Experiment 9. Instructions remained fixed on screen for a period of 10s, after which time words “When you are ready to start the next section, Press Any Key” appeared at the bottom of the screen in 30pt font. When the participant pressed a key, the block began.

The presentation of stimuli and feedback during all blocks was identical to Experiment 4.

Participants were exposed to trials until a criterion of 10 correct in a row was reached. An upper limit of 200 trials was set.

Upon completion of all of the FAST blocks, participants were then asked to complete the Discrimination and Diversity scales.

Results

Function Acquisition Speed Test

Participants completed a practice block, two baseline blocks (one before the test blocks, and one after the test blocks had been completed). Figure 5 below shows the mean trials to criterion for each block completed by the participants.

There was a decrease in the mean trials to criterion between Baseline 1 and Baseline 2. A paired samples t-test showed that this difference was statistically significant ($t = 3.449$, $p = 0.01$). This mirrors the result in Experiment 9, demonstrating that practice effects may indeed be a significant factor influencing critical test results.

A further paired samples t-test demonstrated that the difference between trials to criterion on Consistent blocks (Mean = 30.76, SD = 20.335) and Inconsistent Blocks (Mean = 36.06, SD = 30.696) was not statistically significant ($t = 1.586$, $p = .117$). This indicates that "immigrant" and "cheat" do not have a history of being related at the group level.

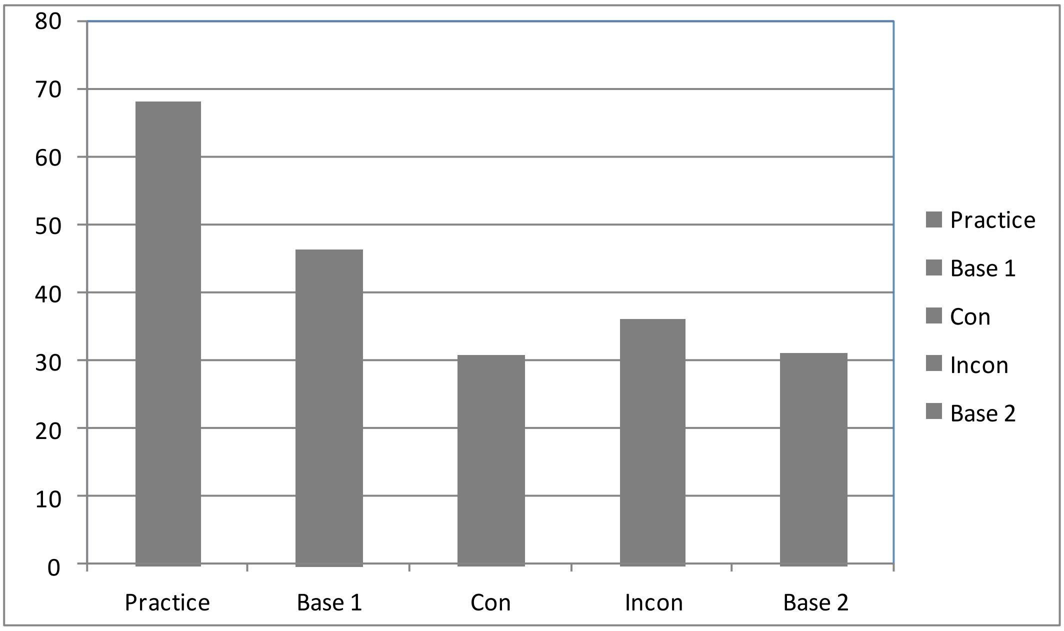


Figure 5: Mean trials to criterion for each block of the FAST.

As in previous experiments, an SoR index was calculated. The mean SoR index was 1.5419 (SD = 8.414). A one sample t-test showed that SoR was not significantly different from zero (t=1.660, p = 0.101) for the entire group. Due to the significant practice effect from Baseline 1 to Baseline 2, the scores were also tested for an order effect with an independent samples t-test. The test showed no effect based on order (t = -.546, p =0.587). In effect, the large practice effect observed across baseline blocks does not account for the absence of an overall FAST effect.

Pen and Paper Measures

Participants also completed pen and paper measures - the Discrimination (DS) and Diversity (DV) Scales. The items were scored such that a score of 1 or 2 indicated negative racial stereotyping, 4 or 5 indicated positive racial stereotyping, and 3 indicated no stereotyping. The mean combined DSDV score was 3.329 (SD = .487), indicating very slight positive racial stereotyping. It is worth nothing that the minimum score on the measure was 2.28, which indicates only mild negative stereotyping. DSDV scores were not significantly correlated with SoR scores (p = .336).

The data were then recoded into “High Prejudice” and “Low Prejudice” Participant’s with a score of <3 were coded as “High Prejudice” and participants with scores greater than 3 were coded as “Low Prejudice”. 59 participants had a mean DSDV score of greater than 3, while 22 had scores of higher than 3. Table 2 below shows the mean SoR for each group. “High Prejudice” participants had a mean SoR of 1.7.750 (SD = 8.819) while “Low Discrimination” participants had a mean SoR of 1.53 (SD = 8.819). An independent samples t-test showed this difference to be non significant (t = .069, p =.945)

It is possible that any effects were mediated by social desirability. The mean social desirability was 15.52 (SD = 4.399) and was normally distributed. A univariate ANOVA was conducted to search for main and interaction effects for social desirability and DSDV scores.

The ANOVA showed that there was a significant effect on SoR for Social Desirability ($F = 6.446$ $p = 0.025$). No effect was observed for DSDV, and no significant interaction effect was observed. To investigate this link further, an independent samples t-test was conducted to compare groups with high or low social desirability. The groups were split on the median (Median Social Desirability = 15). The mean SoR for Low Social Desirability Participants was .41 (SD = 5.476), while the mean SoR for High Social Desirability Participants was 3.99 (SD = 9.983). This difference was significant ($t = 2.015$, $p = .047$). This suggests that individuals with a stronger history of relating *immigrant* and *cheat* show higher social desirability scores than those with little history of relating the terms.

Discussion

The pen and paper explicit measures suggested that anti-immigrant discrimination was low. The lowest combined score on these measures was 2.28, which indicates only mild negative stereotyping (low scores indicate negative stereotyping). This result is not unexpected, given the lack of faith in pen and paper measures which birthed the growth of implicit testing methodologies. However, the FAST results also did not indicate a relation between *immigrant* and *cheat* on a group level.

For a socially sensitive topic like prejudice, social desirability demands are likely to be high. Participants were also tested for social desirability, and those who rated high on social desirability had significantly higher SoR scores – that is, those most concerned about social desirability in the context of an experiment examining prejudice were those who also had the strongest association between the words “immigrant” and “cheat” – that is, those who would be most motivated to hide their prejudice. This is a positive result for the FAST, indicating that the FAST is capable of detecting relations which the participants are motivated to conceal.

However, it was disappointing that no effect was observed overall. The FAST is intended to detect a history of relating two specific stimuli. Unlike the IAT and IRAP, it does not use multiple exemplars to capture an attitude towards a broad category. This level of specificity allows the FAST to detect specific elements of broader attitude, or to exclude them from an analysis. In effect, it is possible that the relation *immigrant = cheat* was simply not prevalent in the sampled group. The effect for social desirability suggests some amount of conscious impression management on the part of some participants, but overall, the sample did not appear to have a history of speaking about immigrants as cheats.

In order to more completely test the FAST, it is necessary to conduct an experiment with an even more socially sensitive relation under examination – one which participants would be motivated to conceal, but that is also likely to be present in the sampled population.

Experiment 11

Measuring naturalistic stimulus relations part 3: Sexual Attraction

Implicit Testing has been of particular interest in the field of forensic psychology, especially with regard to sexual offending. When dealing with sexual offenders, there is a great deal of concern amongst researchers about "faking good" responses during various interrogative procedures. For example, a paedophile population may wish to hide their unchanged sexual attitude towards children following a therapeutic intervention for fear of legal sanctions. Alternatively, they may wish to fake more acceptable attitudes as part of an assessment procedure that may increase chances of parole or other privileges. Therefore, the advent of implicit testing procedures has been welcomed in the field with an increasing number of studies attempting to measure the implicit cognition of offenders, or of at risk segments of the population (Brown et al., 2009; Dawson et al., 2009; Gannon, Wright, Beech, & Williams, 2006; Gray, Brown, MacCulloch, Smith, & Snowden, 2005; Keown, Gannon, & Ward, 2008a; Kamphuis, De Ruiter, Jannssen, & Spiering, 2005; Keown, Gannon, & Ward, 2008b; Malamuth & Brown, 1994; Mihailides, Devilly, & Ward, 2004; Nunes, Firestone, & Baldwin, 2007).

Gray, et al. (2005), Mihailides et al., (2004), and Nunes et al. (2007) have all reported IAT effect differences across child-sex offender and non-offending populations that allowed the two to be distinguished at the group level. In the Gray et al., study, effects bordered on allowing for discrimination at the level of the individual. In another study, Snowden, Wichter, and Gray (2008) used an IAT to discriminate homosexual from heterosexual men successfully. Finally, researchers have also employed modified Stroop procedures to identify sexual interests (see also Ó'Ciardha & Gormley, in press, for a study that employed a pictorial modified Stroop task to explore the sexual interests of child sex offenders).

With these results in mind, the area of sexual interest would appear to be a fruitful one for future application of the FAST procedure. Further, this avenue of research opened up a new opportunity to explore more nuanced uses of the methodology. That sexual interest in teenaged girls is inappropriate for adult males is a strongly socially sanctioned viewpoint. However, there also exist growing concerns in some sectors of society that young women in particular are being sexualised and objectified at a younger and younger age. Furthermore, the concept of the younger woman as being more sexually attractive is common in cultural discourse.

These two conflicting narratives provide an interesting question. At which age point does the stimulus of "young female" move from "non-sexual" to "sexually attractive" for the average male? The FAST procedure measures relations between two specific stimuli. Running a number of FASTs in serial with stimuli that progress along a particular continuum – in this case, age of the female target stimuli – may allow the researcher to map the age point at which an equivalence class between female targets and other sexual stimuli begins. This experiment aims to explore this question using images of young females in three different age brackets. These brackets are defined by "Tanner Age", a widely used system for categorizing a person's physical progression towards adulthood. The other test stimuli used will be sexual images.

Given the problems encountered in Experiment 2, which used picture stimuli intermixed with nonsense syllable stimuli, this experiment also provides an opportunity to rectify those errors and conduct a proper study using the FAST with images and multiple exemplars as stimuli. Given the specificity of the relations probed for by the FAST, these exemplars must necessarily be of a narrow enough range to exemplify a narrow class of stimuli. Further, all stimuli used in the FAST must be of the same form – images, with a small number of exemplars each representing a narrow class of objects. Therefore, where previous FASTs use a single stimulus, the experiment described below will use 3 exemplar images in the place of one stimulus.

Method

Participants

27 participants completed the study. All participants were male. Ages ranged from 19 to 45 (Mean 28.56, SD = 8.135). Participants were recruited from amongst acquaintances of the experimenter.

Apparatus and Stimuli

The FAST blocks were presented to participants on an *Apple Macbook*® laptop computer with a 13-inch monitor (1024 X 768 pixel resolution). Stimulus presentations were controlled by the software package *Psyscope* (Cohen, MacWhinney, Flatt & Provost, 1993) that also recorded all responses.

Stimuli consisted of four real words ("lamp", "rug", "tree" and "board"), thirty-three photographic images taken from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 1999) and nine pseudo-photographic computer generated images from the "Not Real People" images set (NRP; Laws & Gress, 2004) images formed part of the Not Real People Stimulus set (NRP; Laws & Gress, 2004). The NRP is a set of images of males and females that have been computer modified for use in the assessment of and research involving sexual offenders. Each finished image was produced by compiling and morphing three or more images, plus additional modifications such as hair, eye and body colour, poses and clothing. These images were sorted by "Tanner Age", a system which groups individuals based on the particular stage of physical development a person is at. Images of females at Tanner stages 3, 4, and 5 were used, indicating ages of 11.5-13, 13-15, and 15+ respectively.

The experiment also made use of the New Scale of Social Desirability questionnaire (M-C SDS; Crowne & Marlowe, 1960).

Ethics

All participants were presented with and signed a consent form before proceeding to first phase of the experiment (See Appendix 1). Participants were told informally that performance on the task would not allow the researcher to make any individual psychological assessments but may allow for group patterns to be identified. All participants were informed of the true nature of the study after participation and were given the opportunity to withdraw their data at that stage.

General Experimental Procedure

In this experiment, participants completed a practice block, a baseline block, three sets of two FAST test blocks (consistent and inconsistent), and a second baseline block. Each pair of FAST test blocks utilized different test stimulus sets (see below for details). The order in which the three pairs of test blocks were presented was randomized between participants to protect against order effects.

Participants then completed a categorization task to ensure that the stimuli already formed distinct categories and that there was no significant overlap across the various categories of stimuli employed in the main test blocks. Figure 6 below describes the block order and required responses for the FASTs used in this experiment.

Practice Block

The Practice block was identical in all respects to the one presented in Experiments 5 and 6.

Baseline Blocks

Two baseline blocks were presented (one before the first FAST testing, and one after the final FAST block). Each baseline block used four sets of pictorial stimuli taken from the IAPs. Baseline 1 utilised abstract drawings (X1), butterflies(X2), mountains (Y1) and flowers (Y2). Baseline 2 used boats(X3), birds(X4), transport (X5), and rollercoasters (X5). Each picture set consisted of three images. These phases attempted to establish the following relations, respectively; X1-Y1, X2-Y2 and X3-Y3, X4-Y4.

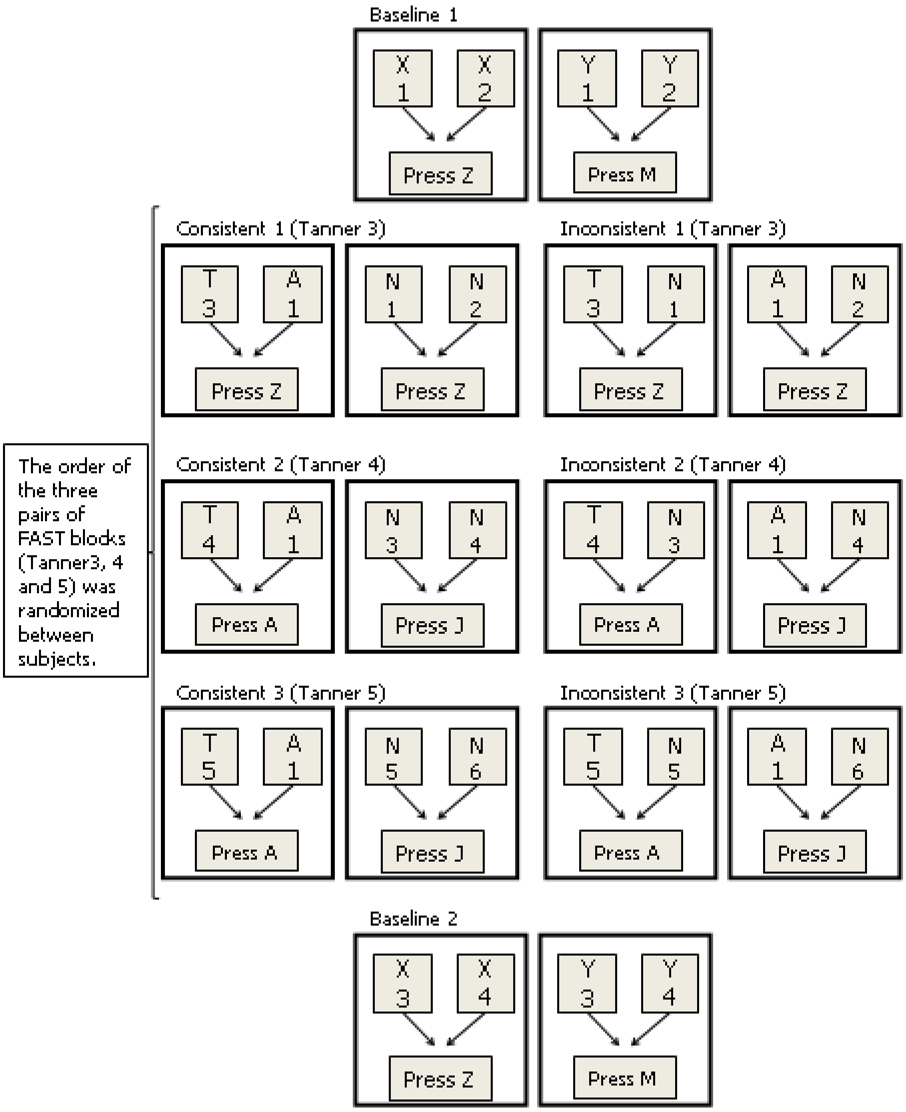


Figure 6. A schematic of the block order and various response functions assigned to stimuli during baseline and FAST blocks. Note that in the current experiment, each alpha numeric designation represents a category of stimuli consisting of 3 images. (see Stimuli and Apparatus)

At the beginning of the Baseline Block, the following instructions were presented:

IN THE FOLLOWING SECTION

A single picture will appear on screen

For Example

Each time you see the word you must press either

The “**Z**” key or the “**M**” key

After you press a key, the computer will tell you if you were

CORRECT or WRONG

If you don’t press anything within 2 seconds of seeing a word, the computer will count this as a WRONG answer and move on to the next word

Your task is to learn which key **(Z or M)** to press for each of the words. This section will continue until you can make 10 CORRECT RESPONSES in a row WITHOUT ANY MISTAKES

Instructions remained fixed on screen for a period of 10s, after which time words “When you are ready to start the next section, Press Any Key” appeared at the bottom of the screen in 30pt font. When the participant pressed a key, the block began.

All trials were presented on the computer screen with a white background. A trial began with the presentation of one of the twelve picture stimuli in the centre of the screen.

Feedback, trial order, response windows, and maximum trial limit were identical to the *General FAST Block Procedure* used in Experiment 4

After the first Baseline Block was completed by the participant, a screen appeared with the following text:

Well done! You have completed this section of the experiment.

Please take a short break to rest your eyes from the computer screen.

The instructions for the next section will be displayed shortly.

This text remained on screen for 30s, after which the same instructions as above were displayed, for the same duration and before.

Each test block utilized four stimulus sets, each consisting of three images (See Appendix 2). The stimulus sets of interest were a set of three erotic images of couple (T1) and three images of young girls from the NRP stimulus set (T2). The first set of FAST test blocks utilized images of girls at Tanner Age 3, the second pair of test blocks utilized girls at Tanner Age 4, and the third pair of test blocks utilized girls at Tanner Age 5The neutral stimuli (N1 and N2) were also images of people, in order to avoid confounds associations. The first set of neutral stimuli was three images of skiers (N1). The second of these were 3 images of people engaged in group gatherings (N2).

Participants were exposed to three pairs of FAST test blocks. The first of the pair (consistent) established two functional response classes (T1-T2 and N1-N2) that would be consistent with relations predicted given a history of child-sex relational responding on the part of the participant. The second test block established two functional response classes that were inconsistent with those relations (T1-N1 and T1-N2). The second and third pairs of test blocks utilised different T2 stimuli (see above).

As in Experiment 9, it is impossible to determine *a priori* which of these two sets of relations are actually "consistent" with the participants' history – the naming of these blocks as "consistent" or "inconsistent" is merely an experimental convenience. The order of block presentation was randomised across participants.

After each FAST Block was completed by the participant, a screen appeared with the following text:

Well done! You have completed this section of the experiment.

Please take a short break to rest your eyes from the computer screen.

The instructions for the next section will be displayed shortly.

This text remained on screen for 30s, after which the following instructions were displayed:

The instructions presented at the start of these blocks were identical to those used in the baseline and test blocks for Experiment 9. Instructions remained fixed on screen for a period of 10s, after which time words "When you are ready to start the next section, Press Any Key" appeared at the bottom of the screen in 30pt font. When the participant pressed a key, the next block began.

Feedback, trial order, response windows, criterion, and maximum trial limit were identical to the *General FAST Block Procedure* used in Experiment 4

In summary, participants were exposed to three test blocks, each measuring the strength of relation between images of children at three different stages of puberty and erotic images. These blocks were preceded by and followed by a baseline block.

Categorization test

In this phase, participants were exposed to a categorization test. Stimuli consisted of all images employed in the FAST test blocks (all images from the consistent and inconsistent blocks). This test served to ensure that the categories and the stimuli representing them were clearly discriminable as discrete stimuli and non-overlapping categories. Prior to the onset of the test, participants were presented with the following on-screen instructions:

Just one more section to go!

This is a quick categorization task

Press 1 if the picture is a GIRL picture

Press 2 if the picture is a SKIER picture

Press 3 if the picture is a SEXUAL picture

Press 4 if the picture is a GROUP picture

Press any key when you are ready to begin

Once the participant pressed a key to begin, the last line of instructions disappeared, and a single photographic image was displayed below the instructions which remained. Stimuli were presented once each in a random order. Once a participant responded with one of the four instructed key presses, the image cleared and the next picture was displayed. No feedback was presented during this phase.

Once the participant had viewed each picture once and responded to each, the screen cleared and the following text was displayed:

Congratulations!

You have completed the final section of the experiment

Please contact the experimenter

Thank you very much for your time

Marlowe-Crowne Social Desirability Scale

After completion of the computer task, the New Scale of Social Desirability questionnaire (M-C SDS; Crowne & Marlowe, 1960) was administered to each participant.

Results

Baseline Blocks. Each participant completed two baseline blocks. The mean trials to criterion for the first Baseline Block was 120.19 (SD = 73.195). The mean trials to criterion for the second Baseline Block was 76.67 (SD = 72.862). A paired sample t-test showed that the difference from Baseline 1 to Baseline 2 was significant (t = 2.321, p = -0.028). As noted earlier, the order of the intervening blocks (with regard to the "Tanner Age" variable, not the order of consistent and inconsistent test blocks within a pair) was randomised to protect against practice effects which was clearly required given evidence of this significant practice effect.

FAST Block 1 (Tanner Age 3). In the FAST procedures that follow, the block in which the female stimuli and sexual stimuli shared a response are referred to as the consistent block. . In the first FAST block (in which female stimuli were of Tanner Age 3), the mean trials-to-criterion for the consistent block was 47.15 (SD = 40.917) and the mean trials-to-criterion for the inconsistent block was 29.67 (SD = 25.046). The mean difference between trials-to-criterion was -17.48 (SD = 43.345).

SoR scores were calculated for each participant, using the mean of the two baseline phases that preceded and followed the FAST test blocks. The mean SoR was -4.832 (SD = 11.115). Of the 27 participants, only 5 participants had an SoR of greater than zero, indicating a widespread tendency to not easily acquire common response functions for sexual images and images of pubertal girls. The mean positive SoR was 8.33 (SD = 7.709). This mean was substantially increased by the SoR score of S11 (SoR = 21.96). The mean negative SoR was -7.824. (SD = 8.619).

A t-test indicated that, on the group level, the SoRs were significantly different from zero. (t = -2.259, p = .033),

FAST Block 2 (Tanner Age 4). In the second FAST block (in which female stimuli were of Tanner Age 4), the mean trials-to-criterion for the consistent block was 35.19 (SD = 22.76) and the mean trials-to-criterion for the inconsistent block was 48.52 (SD = 48.578). The mean difference between trials-to-criterion was 13.33 (SD = 54.570).

SoR scores were calculated for each participant. The mean SoR was 3.306 (SD = 13.21). Of the twenty-seven participants, fourteen participants had an SoR of greater than zero. The mean positive SoR was 11.995 (SD = 12.151), while the mean negative SoR was -6.051 (SD = 5.939).

A t-test was conducted. These SoR scores were not significantly different from zero (t = 1.3)

FAST Block 3 (Tanner Age 5). In the third FAST block (in which female stimuli were of Tanner Age 5), the mean trials-to-criterion for the consistent block was 40.40 (SD = 40.816) and the mean trials-to-criterion for the inconsistent block was 63.25 (SD = 66.767).

SoR scores were calculated for each participant. The mean SoR was 1.3649 (SD = 15.606). Of the twenty seven participants, sixteen participants had an SoR of greater than zero, and ten participants had an SoR of less than zero. The mean positive SoR was 8.976 (SD = 12.398), while the mean negative SoR was -11.573. (SD = 11.612). One participant had an SoR of exactly zero.

A t-test was conducted. These SoR scores were not significantly different from zero (t = .454). Figure 7 below illustrates the mean SoR scores for each pair of FAST test blocks and thus the three varying Tanner Ages of the female stimuli.

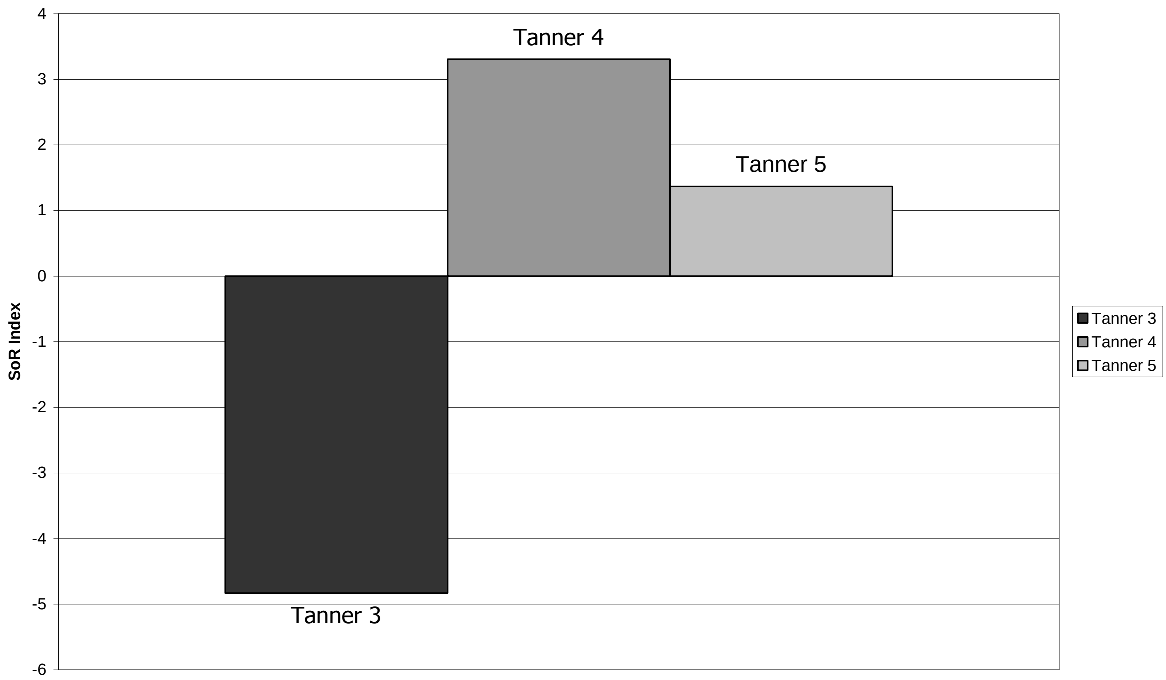


Figure 7: *Mean SoR Indices for each pair of FAST Test Blocks.*

Summary Participants completed three consecutive pairs of FAST test blocks. In each FAST test block, the Tanner Age of the female stimuli was varied. Tanner 3 indicates ages 10.5-13, Tanner 4 indicates ages 13-15, and Tanner 5 indicates ages 15+. Participants found it very difficult to complete FAST blocks in which Tanner 3 female images required the same response as sexual images. The SoR scores on this FAST resulted in significantly negative SoR indices. This suggests that pairing those two stimuli was strongly opposed by the participants' verbal history. By contrast, when the Tanner Ages of the female images increased to 4 and 5, the results were highly variable within the participant group, and did not yield a statistical difference from zero on the group level.

Discussion

The results of this experiment demonstrated that the FAST is capable of being used in serial to identify the boundaries of a relational network in a group of participants. As an overarching class, “children” are not likely to be related to “sexual” in the normal population. In the condition where exemplars were in their early teens, the participants demonstrated a strong resistance to forming similar functional responses to those images and sexual images, indicating a strongly incompatible relational history that impeded their FAST performance. However, with images of females in the mid and late teens, the acquisition of the functional responses required to complete the consistent FAST block was less impeded, producing no significant differences in learning rates between inconsistent and consistent blocks on a group level. However, considering the high standard deviations observed, and the high mean negative and mean positive SoRs for those groups, this lack of significance may be more because of within group variation that a group tendency towards zero difference.

As well as identifying a history on the part of a normal population experimental group, the current experiment also highlights an important feature of the FAST format – that is it appears to be highly effective at identifying stimuli which share a history of being mutually exclusive. The mean SOR in the Tanner 3 condition was the strongest observed in any of the experiments in this chapter (-4.832), and this was substantially affected by a large positive outlier. This indicates that completing a block which requires common responses for incompatible stimuli is substantially more difficult than completing a block with similar responses for compatible stimuli. This has implications for the stimulus pairings which are most likely to produce strong and reliable FAST effects. In summary, the FAST may be more suitable for measuring incompatible stimulus relations than compatible ones. This makes sense even in numerical terms insofar as a floor effect is quickly reached when the acquisition of common response functions by otherwise neutral stimulus pairs is accelerated by a prior history of shared function. On the other hand, such a learning rate can be greatly slowed by a history that is incompatible with current contingencies. This should be borne in mind by future researchers.

In this experiment, there was also a statistically significant practice effect observed. The order in which the FAST test blocks were presented was randomised across participants, but it must be acknowledge that over multiple presentations of the FAST, that practice effects will occur across Baseline phases. This is the primary reason that the mean of Baseline performances is taken for the purposes of calculating SoR indices.

General Discussion

Each experiment in the current chapter probed for naturally occurring stimulus relations. Under the Relational Frame Theory account of attitudes, attitudes can be described as networks of stimulus relations reinforced and maintained through interactions the socioverbal community. The specific relations targeted in these experiments were stimulus relations that were hypothesised to be key parts of the larger networks which constitute attitudes. Data from pen and paper measures targeting the explicit attitude construct were gathered, and social desirability questionnaires were also administered in order to compare FAST performance to pen and paper measures.

Experiment 9 targeted the relation spiders = disgust in eighty-eight participants. The relationship between disgust and fear in the "arachnophobia" attitude construct has been a question that researchers in the field have previously attempted to tackle using implicit measures, but the results have failed to clarify the question. The results of Experiment 9 showed strong positive FAST effect for the *spider-disgust* relation, indicating that "spiders are disgusting" is a common relation reinforced by the verbal community. However, the sample had a low rate of spider fear according to the pen and paper measures (FSQ; Szymanski and O'Donohue, 1995), suggesting that this relation is independent of explicit spider fear. These results were interpreted as evidence that "spiders are disgusting" is a verbal relation that most participants were highly fluent in, allowing them to easily emit the same responses to both stimuli during the consistent blocks of the FAST. The Gavin et al. (2008) model interprets IAT performance in terms of an individual's verbal history and practices, in accordance with the RFT model of attitudes. This same interpretation is applicable to FAST performances.

In Experiment 10, the two stimuli under investigated were the words *immigran*t and cheat.Eighty-seven participants completed the FAST and pen and paper measures. This relation was chosen on the basis that ethnic prejudice in Ireland is not so much based on race as based on an immigrant/native divide. Explicit measures (the Discrimination and Diversity Scale; Wittenbrink et al., 1997) indicated a low level of explicit prejudice in the sample, and this result was corroborated by the FAST results, which showed no group level effect for the relation *immigrant = cheat.* It is worthy of note that participants who rated high on Social Desirability (Crowne & Marlowe, 1960) also tended to have higher SoR indices than those who scored low on that scale. As such, it was concluded that *immigrant = cheat* was not a relation in which the sample was fluent, but that those individuals who were more fluent also felt more pressure to emit social desirable responses in the testing context.

Experiment 11 targeted relations indicative of sexual interest in young females by using erotic image stimuli and images of young females. Twenty-seven male participants were tested Unlike the previous two experiments, the stimuli used in this FAST were image stimuli, and sets of 3 exemplar images where used in place of a single word stimulus. Further, 3 pairs of FAST test blocks were completed by each participant, with the Tanner Age of the exemplar female images being varied in each run. Tanner Age is a means of categorizing the physical maturity of a person. Experiment 11 utilised images with Tanner Numbers ranging from 3 to 5, representing a range of ages from 10.5 - 15+. The results of this experiment indicated that the FAST is highly sensitive to stimuli which participate in a frame of opposition – the FAST testing for relations between Tanner 3 females (Ages 10.5 – 13) and sexual images showed the highest recorded negative mean SoR in this chapter. The older female images showed a great deal of within group variability, but no group level difference from zero. This suggested that while females aged younger than 13 were clearly and strongly non-sexual, this relation of opposition was not present in the group for teenaged girls in general. This reflects the equivocal messages in society with regard to the sexualisation of young women.

Taken together, the results reported in this chapter are highly informative as to the particular strengths and weaknesses of the FAST procedure in an applied context. The

FAST is capable of detecting relations which are strongly endorsed by the socioverbal community, with which participants can be expected to be highly fluent. “Spiders are disgusting” and “Pre-teenaged females are not sexual” are uncontroversial statements endorsed by the culture, and are verbal relations that will have been spoken and heard often enough that individuals will be highly fluent with them. It is of particular note that the FAST was extremely sensitive to the Tanner 3 FAST which required responding equivalently to sexual images and pre-pubescent female images. This finding has implications for stimulus selection in future experiments, as the FAST may be maximally effective when the Test Stimuli are opposed rather than related. The Watt et al. (1991) procedure likewise focused on socially incompatible stimulus relations (Catholic names and Protestant symbols). The results of Experiments 6 and 8 aid in the interpretation of the results of Experiment 10. *Spiders=Disgust* is a culturally accepted relation, while *Pre-teen females=sexual* is very much the opposite. *Immigrant-Cheat* may not have enjoyed either status in the sampled group. University campuses are not usually places where explicit ethnic prejudice is welcomed, and the targeted relation was very specific. However, the production “immigrants are cheats” may have been common enough in participants’ history that it was neither a culturally endorsed relation, nor a culturally opposed relation.

The extreme specificity of the FAST is both a limitation and strength. If specific relations (i.e. *immigrant-cheat*) are used to infer a more complex attitude (i.e. ethnic prejudice) without proper consideration, then it may seem that the FAST is a weak measure of such attitudes. However, the FAST is not, in and of itself, a measure of attitudes – in the sense that attitudes are complex relational networks. The FAST measures specific stimulus relations, which are the constituent elements of an attitude, not the attitude entire. Future researchers must take care to keep this in mind when designing studies using the FAST methodology.

The same specificity can also be leveraged to create clearer pictures of the particular stimulus relations which underpin an attitude. As noted in Experiment 9, untangling the different contributions of disgust and fear-of-harm in spider phobic individuals is an important research target, and the dyadic nature of other implicit measures has been problematic for researchers when it came to interpreting their results. As demonstrated in Experiment 11, running a number of FASTs in serial can be used to explore the specific details of a set of stimulus relations.

This type of use of the FAST can be thought of as mapping the limits and shape of stimulus classes. Consider a thought experiment in which Experiment 11 was repeated. Instead of limiting the female image stimuli to a narrow range of ages, imagine running 60 FASTs in serial, with images ranging in age from 5 to 65 years old. The graph that would be produced would be similar to Fig. 7, but with many more data points. That graph of mean SoRs, ranging across all ages, would yield a picture of cultural attitudes with regard to sexually appealing females. It would show a band of ages in which *female=sexual* was strongly rejected. There would be a band in which the average would hover around zero, representing the ages around which there are mixed cultural messages. We would then see a band of strong positive SoRs, followed by another dip as the exemplars aged.

Similarly, with regard to the ethnic prejudice FAST in Experiment 10, words similar to “cheat” could be used as Test Stimuli in a series of FASTS. Words such as liar ”,“ fraud and “thief” may show a relation that was not present in a FAST which targets a single relation. This kind of fine analysis is most appropriate with large sample sizes, as the relations of interest are those which are endorsed across a culture or subculture. The IRAP and the IAT, given the dyadic nature of the stimuli which must be used, would struggle to break a complex attitude down into its constituent stimulus relations. Further, a single FAST is substantially quicker to administer (conservatively, 4 minutes per block), so multiple FASTs in serial is more feasible than multiple IRAPs, merely from a pragmatic point of view.

Chapter 5

Discussion and Conclusions

In the remainder of the final chapter of this thesis, I will discuss the experimental findings from each of the previous chapters, as well as the implications of these findings. In doing so, we will chart the early development of this novel test format, and discuss the remaining challenges that must be confronted in future research. Issues relating to wider research on implicit testing and the nature of unconscious attitudes will be discussed to some extent. Alternative experimental preparations emerging from this research will be considered and explored, and possible applications in both basic research and applied setting will be discussed in detail.

In Chapter 2, a series of experiments were conducted to establish proof-of-concept for the FAST methodology by training a set of known relations and deploying the FAST procedure to detect them. The experiments detailed in Chapter 2 each used a similar experimental format. Participants were exposed to a pre-experimental training procedure to establish a known history of relations between stimuli, which would then be probed for using a FAST procedure. The specific relations trained varied from experiment to experiment (see below) and the precise FAST procedure used also varied as small methodological changes were made to fine tune the effectiveness of the procedure. This investigative strategy allowed the researcher to avoid the difficulties involved in stimulus selection that occur when using real-world stimuli. Naturally occurring stimulus relations, while often strongly established and well rehearsed on the part of the participant, are difficult to predict pre-experimentally. The ambiguities created by starting with testing of naturally occurring histories would cast doubt on the results. The question would arise; is the FAST be failing to detect a relation because the relation is not established for that participant sample, or because the procedure is flawed? The creation of laboratory controlled histories allows us to make clearer predictions and better assessments of the FAST results.

In Experiment 1, twenty-one participants were exposed to a simple matching to sample procedure to establish the relations A1-B1 and A2-B2. Next, the novel FAST test format was used, with A1 and B1 being the two test stimuli. Given that the relations being tested for had been established using a well-established training procedure, it was expected that the FAST would capably detect the pre-experimentally trained relations. Thirteen out of the eighteen participants showed the expected “FAST effect”. That is, they took longer to complete inconsistent blocks that consistent blocks. This was a promising result for the novel procedure, providing a foundational proof-of-concept for the FAST. Subsequent experiments were designed to expand the range of relations tested by the FAST.

Experiment 2 was an extension of Experiment 1. Eleven articipants were exposed to a function training protocol which established an erotic function for A1 and an aversive function for A2. In the FAST block which followed, the test stimuli were A1 and a trio of the erotic images. In this experiment, participants were exposed to three runs of the FAST procedure in succession. In the first run, only two out of ten participants showed the expected FAST effect. Eight of ten participants showed the expected effect in the second run, and six out of ten showed the expected effect in the expected direction in the final run. However, the effect was not statistically significant at the group level in the final run. This was the first attempt to run a FAST procedure using picture stimuli, and it was the only FAST in which a mixture of pictures and words were utilized. This experiment also flagged a concern that would recur in future experiments – that of the FAST’s stability across multiple runs and the attendant issues with block order.

In Experiment 3, twenty-two participants completed a matching-to-sample procedure that established two 3-member equivalence classes (A1-B1-C1 and A2-B2-C2) according to a one-to-many (sample-as-node) protocol. Once again, three letter nonsense syllables were used as stimuli. Three runs of the FAST then probed for derived A1-C1 relations. For a testing format which aspires to being able to detect socially relevant verbal histories, the ability to detect derived relations is of primary importance, as derived relational responding is thought to be the underlying process for language according to many modern behaviour analysts (e.g., Sidman, 1994; Hayes, Barnes-Holmes & Roche, 2001). In contrast to Experiment 2, all three runs of the FAST showed a statistically significant FAST effect in the expected direction for participants who passed the equivalence testing phase prior to the FAST.

Given the importance of derived relations in language processes, the remainder of the experiments in Chapter 2 focused on extending Experiment 3 and further testing the ability of the procedure to detect derived equivalence relations. While experiments 1 and 3 showed encouraging results, there was still some concern remaining with regard to the number of participants who failed equivalence testing in Experiment 3 and in the poor results for Experiment 2. It was unclear at that point as to whether the result was due to problems with the FAST format, or whether the inconsistent results were due to poor stimulus control in the pre-experimental training phases. As such, in Experiment 4, a number of changes were made to improve stimulus control.

First, the matching to sample training that was completed by participants was designed to be extremely robust. Participants completed nine phases of matching to sample training, over three different sessions which spanned a week. The criteria for completion of each training and testing phase became more stringent as the phases progressed. Participants had to complete blocks with increasing lengths, longer delays between observation responses and the appearance of response options, and a graduated reduction in reinforcement contingencies. All of these changes were put into place to maximise the probability of equivalence performance emerging at the end of the training. All participants successfully passed the equivalence testing block in a single cycle.

Furthermore, the instructions presented to participants at the onset of the FAST were altered to make them as clear as possible. A practice block was included to familiarize participants with the FAST task before testing began. A further addition was made to the FAST procedure – the inclusion of a counter which showed the participant the number of correct answers in a row they had achieved, and thus how close they were to completion of the block. This also served to signal the punitive consequences of an incorrect answer, or of failing to respond within the response window.

Due to the nature of the training procedure, Experiment 4 was conducted as a repeated single participant design. In the FAST which followed the extensive training, four out of five participants showed a strong FAST effect, with highly positive FAST effects. This is perhaps the most important experimental result in this thesis, providing strong evidence that the FAST reliably detects a history of derived relational responding when those relations are well rehearsed and extensively trained.

Experiment 5, the final experiment in Chapter 1, investigated the FAST's potential to detect untested derived relations. An equivalence testing procedure requires a participant to tact the derived relations, and as such, makes the implicit explicit. However, when considering the role of derived relations in attitude formation, in cases where the implicit attitude is not available to conscious awareness, untacted derived relations may play a role. In Experiment 5, thirty eight participants were split into two experimental groups. The first group completed a linear matching to sample training procedure, followed by a FAST, followed by an equivalence testing procedure. The second group completed the FAST after the training and testing phases. The FAST did not prove sufficiently sensitive to detect any derived relations before equivalence testing. However, the FAST did reliably detected relations which when both training and testing phases had been completed prior to administration of the FAST.

The experiments detailed in chapter 1 show both the limits and the potential of the FAST procedure. The FAST is capable of detecting short experimental histories of equivalence responding (Experiments 1, 3, and 5). The positive results in these occurred in the context of only recently-established stimulus relations. Even with the extensive training used in Experiment 4, these relations cannot hope to approach the level of rehearsal and reinforcement that the real world relations underlying attitudes have in a socio-verbal community. Nevertheless, the fact that the FAST detected these relations in several experiments provides a solid foundation for further research, because the real world relations for which the experimentally trained relations are an analogue will be significantly more rehearsed in the participants' behavioural history.

The FAST failed to detect histories in which the relevant stimulus relations were not previously derived, or were untacted (Experiment 5). This is an important finding with considerable implications for the implicit tests that harness the same core process as the FAST (i.e., the IAT and the IRAP). More specifically, the relational frame theory account of attitudes describes an attitude as *a network of derived and explicitly reinforced stimulus relations according to which the functions of events are transformed* (Barnes & Roche 1997; Dixon et al., 2003; Grey & Barnes, 1996). According to this view, attitudes may consist of relations not yet derived explicitly, but merely implied. As an example (cf. Dixon et al., 2003), the parents of a child may have regularly referred to people of *Irish* origin as *drunkards* while speaking to their child. In other contexts, the parents, or other individuals, may have referred regularly to *drunkards* as *ignorant.* These contingencies parallel a linear

(A-B-C) stimulus matching sequence in which a derived relation between the terms *Irish* and *ignorant* might be expected to emerge. While there may have been no occasion on which an individual was instructed that the Irish are ignorant, Dixon et al. (2003) suggest that this history would be sufficient to produce a prejudice against Irish people. Experiment 5 sought to explore the ability of the FAST to detect histories of untacted derived relations, and failed to do so.

However, the consistent ability of the FAST to detect equivalence relations that have been both trained and tested, and to detect naturally occurring verbal relations which have been well rehearsed (see Experiments 9 and 11), is in stark contrast to the apparent insensitivity of the FAST to untacted relations. Of course, as yet untacted relations may still emerge at some point in the future, and the baseline relations established by the verbal community may still make certain prejudices or other biases more likely, even where they are not explicitly taught. However, the current results suggest that relations must nevertheless be explicitly derived before they can be detected by implicit-style methodologies.

The research question regarding the detection of as yet untacted stimulus relations (e.g., unconscious stimulus relations) is one which the FAST may be uniquely positioned to address. It is, *prima facie*, impossible to determine whether an equivalence class has formed without requiring a participant to discriminate directly that relation in an equivalence test. However, the FAST procedure does not require that the two stimuli be presented together on one trial. The FAST method was designed to harness the behavioural processes identified by Watt et al. but using a procedure that would be quick and easy to administer, while at the same time retaining its procedural subtlety, and even improving upon it. More specifically, the FAST was designed to allow researchers to quickly calculate the likelihood that any two stimuli have been related in the history of a participant, while at the same never requiring the participant to directly tact the relation under examination. The IAT, IRAP, and Watt et al

(1991) procedure all require the participant to respond by directly relating the items under investigation in each trial. The FAST requires the participants to learn specific functional responses to a single stimulus per trial, and as such is more indirect in its investigation. Finally, unlike the IAT and IRAP, no contingency specifying rules are presented for responding (i.e., responding is shaped using corrective feedback) in the FAST, and thus the test-taker is never presented with the two verbal stimuli of interest simultaneously during the test or at any other time. As such, the FAST methodology can in principle probe for the emergence of a derived relation without it ever having to be directly tacted.

While Experiment 5 was unsuccessful in detecting untacted relations, this only confirms that untacted relations do not appear to impact upon future FAST behaviour where only a short history of relational training has been provided. Future experimental work can utilise more extensive training procedures, similar to those used in Experiment 4. Naturally occurring verbal relations undergo staggering amounts of rehearsal and reinforcement in a normal socioverbal environment, so a great deal of laboratory training would be needed to generate relations nearly as strong.

In addition to the advantages offered by the FAST's *procedural implicitness* (De Houwer, 2006), the FAST might also allow for the concurrent assessment of a matching response repertoire even as the very repertoire is emerging. This might be useful for assessing the long learning curves involved in producing an experimental analogue of natural language in relational training experiments or teaching programmes. A slight change in a participant's FAST performance may be important in determining the precise conditions under which derived relations can affect behaviour prior to being explicitly tacted. It is possible that even before forming stable patterns of responding, certain histories may nevertheless be stable enough to interfere with the acquisition of other functional response classes (e.g., responding to a variety of white and non-white individuals in an identical and courteous way).

If it is possible to use the FAST to detect changes in “relational strength” as repertoires of derived relating emerge, it would suggest that social histories may be capable of interfering with further learning by virtue of implied stimulus relations of which individuals themselves are not even aware. For example, a history in which the word “Black” is related to the word “African’, which is in turn related to the word “lazy”, may impede learning an association between a black person’s face and positive words describing the nice sweater they are wearing. In other words, even following attempts to teach an individual something positive regarding a particular Black person, the acquisition of the required stimulus relations (e.g., *Black person- nice*) may be slower than if the relevant stimuli did not participate in any conflicting implied (underived) relations (e.g., *Black person – not nice*). Thus, an individual’s history may create resistance to such learning, perhaps without the individual ever being able to report why that is the case. Other psychologists may refer to this as an “unconscious process”, but given the current functional-analytic approach, it is possible to keep the process of learning interference overt and open to direct experimental analysis.

It must be emphasized that implicit testing emerged within a social-cognitive paradigm, and research in this field takes a particular interest in unconscious over conscious stimulus associations (see Hughes, Barnes-Holmes & De Houwer, 2012 for a behaviour-analytic review of this issue). However, social-cognitive researchers do not have a paradigm within which they can generate implied relations under laboratory conditions, and as such, researchers must rely on presumptive implicit attitudes such as “flowers-good” and “insects-bad” to establish the effectiveness of their measures. The functional-analytic underpinnings of the FAST and the behavioural tradition which inform its processes give a particular advantage to using the FAST in this line of enquiry, as we are in possession of the methodology which allows one to generate derived relations. In effect, a derived relation may simply be regarded as one that is implied…or implicit. This simple change in definition of the term “unconscious” moves the term from the mentalistic to the functional domain, and opens up whole new avenues for exploring the processes underlying implicit tests, and building better ones whose processes are well defined and in which the behavioural phenomena of interest are equally well understood.

The functional approach can allow us to establish the conditions of the emergence and influence of implicit attitudes, providing a functional underpinning to a research area which has heretofore lacked a robust theoretical basis for its phenomenon of interest. If the relational frame theory account of attitudes is to make a strong impact on social psychology, it must first build on the strengths of its origins and establish the strength of its claims with careful basic research.

These future directions aside, the progression of results throughout Chapter 2 represents a slow extension of the proof-of-concept for the FAST procedure. Beginning with the simplest relations between stimuli and extending into derived relations of varying strength. With the exception of Experiment 2, the FAST proved to be broadly effective in detecting the pre-trained relations, and as such, likely to be similarly proficient at detecting naturally occurring stimulus relations.

In Chapter 3, the researcher turned his attention to the effect of small procedural modifications on the FAST procedures. Each experiment was designed with a similar procedure. Participants completed a matching-to-sample procedure to establish the relations A1-B1 and A2-B2 between pairs of three letter nonsense syllables. Thereafter, participants completed a series of FAST procedures. Three pairs of test blocks were completed, with the test stimuli being the A1 and B1 stimuli from the matching to sample pre-training. Four baseline blocks were used, one prior to each of the test blocks, and one after each pair of test blocks. In each experiment, a single procedural change was tested – response windows, response criterion, and practice blocks. Participants were divided into a number of experimental groups, each of which completed a FAST with a variation on the modification of interest.

Experiment 6 measured the effect of varied response windows on FAST performance for fourty-four participants. Group 1 completed FAST blocks with a 1000ms response window. Groups 2, 3, and 4 completed FAST blocks in which the response windows were 2000ms, 3000ms, and 4000ms respectively. The results of this experiment were difficult to interpret due to instability in the predicted FAST effect. However, in the third pair of FAST test blocks, Groups 1, 2 and 3 showed a significant FAST effect, while Group 4 failed to show any effect.

This result suggested that shorter response windows were important for the FAST's ability to detect participant histories. The FAST procedure is leveraged on the idea that it takes longer to achieve fluent responding on inconsistent blocks in comparison to consistent blocks. Fluency is composed of both response accuracy and speed. As such, a smaller response window represents a higher fluency criterion. It was to be expected that higher fluency requirements would increase differences between easy and difficult blocks (i.e., consistent and inconsistent) as more stringent fluency requirements would have a disproportionate effect when the task is more difficult. However, the evidence also suggests that a 1000ms response window may be too stringent, as a high number of participants failed to complete FAST blocks under these conditions.

Experiment 7 investigated the effect of including a variable number of practice blocks prior to the first baseline block with a sample of twenty three participants. A practice block was identical in design to the baseline blocks, with the exception that the instructions presented flagged them as practice. The three experimental groups received three, five, or seven practice blocks before beginning the same FAST sequence used in the other experiments in Chapter 3. Participants did not display a linear decrease in trials-to-criterion across practice blocks – instead, there was a steady increase in trial requirements, followed by a sudden drop at the fifth practice block. Further, none of the groups showed a significant FAST effect at any point during the three test runs. This suggests that extensive practice may have a deleterious effect on the FAST as participants develop an effective response strategy which overrides the pre-experimental training.

Experiment 8 varied the block completion criteria.There were thirty participants in total. Participants in Group 1 were allowed three errors in any run of ten FAST trials (7/10), while participants in Groups 2 and 3 required eight out of ten and nine out of ten correct responses respectively. Similarly to Experiment 6, it was found that a too lax fluency requirement (7/10) rendered the FAST ineffective. Participants showed the most consistent FAST effects when there was an 8/10 block completion criterion. Again, small sample size and unstable results on the FAST blocks made interpretation of the effect difficult.

Many of the changes suggested by the experiments described in Chapter 3 were incorporated into the FASTs which followed. At this point in the development of the FAST, the "standard" procedure had the following features. Participants complete a practice block, a baseline block, a consistent and inconsistent block (test blocks), followed by a final baseline block. On each trial, there is a response window of 2000ms. The criterion for completion of a block of trials is ten correct responses in a row.

An ongoing challenge in the design of the FAST is the tension between the need to require a high level of fluent responding, and the need to minimize participant attrition. Setting the criterion at 10 correct responses without error, while also reducing the response window from its initial 3000ms level represents a significant increase in task difficulty. In order to mitigate this change, the instructions presented to participants were clarified and reformatted to make the response contingencies clear to the participants. Further, in order to increase the salience of the requirements, a counter was included which displayed the number of consecutive correct answers to the participant. The inclusion of this counter, however, made the implementation of a less stringent response criterion impossible to realise without further increasing the complexity of each individual trial presentation. It was judged that the increase in contingency salience provided by the counter outweighed the suggested benefits of a less stringent accuracy criterion.

In Chapter 4, the research turned to utilizing real world stimuli. In each experiment in Chapter 4, participants completed FAST procedures aimed at detecting natural histories of verbal behaviour. In each experiment, a different relation was targeted. In Experiment 9, the FAST probed for a relation between the words “spider” and “disgust” in a random sample from the general population. Experiment 10 targeted the relations “immigrant” and “cheat”. Experiment 8 departed from using real words as words as stimuli, instead using images of teenaged females to probe for relations between sexual images and images of teenaged females of different ages. The experiments in this chapter took on a research strategy aimed at further exploring the limitations and strengths of the FAST procedure in different contexts – socially sensitive topics (ethnic prejudice), culturally ubiquitous verbal relations (spiders are disgusting), and concepts which are rejected by cultural narratives (young girls are sexual).

Experiment 9 utilised the words “spiders” and “disgust” as test stimuli to probe for an equivalence relations between these words in a sample of Irish college undergraduates. The neutral, unrelated stimuli used in both the baseline blocks and in the tests blocks were also real words. The FAST was administered to eighty-eight participants, along with explicit pen and paper measures – the Fear of Spiders questionnaire (Szymanski and O’Donohue 1995) and the Social Desirability questionnaire (M-C SDS; Crowne & Marlowe, 1960). Participants showed low scores on Fear of Spiders overall. As a result of this, when participants were median-split into high and low spider groups, the FAST failed to detect difference between groups. However, the FAST detected a relation between the words “spider” and “disgust” in the sample as a whole. This result shows that the FAST is capable of detecting verbal relations that are common in the vernacular (“spiders are disgusting”) which are familiar and well rehearsed even in people who are not fearful of spiders, allowing them to easily emit the same responses to both stimuli during the consistent blocks of the FAST. This is fitting with the Gavin et al. (2008) behavioural model which interprets IAT performance in terms of an individual’s verbal history and practices. This same interpretation is applicable to FAST performances.

Experiment 10 used a similar procedure to Experiment 9 with the words “immigrant” and “cheat” used as test stimuli, in order to probe for prejudicial attitudes in a sample of Irish university undergraduates. While the relation probed for in experiment 6 was uncontroversial, overt expression of the attitude “immigrants are cheats” would be considered socially unacceptable by many people in the culture. This kind of socially sensitive relation is precisely the type of relation that implicit testing was developed to probe for. The explicit measures used in Experiment 10 were the Social Desirability Questionnaire and the Discrimination and Diversity Scale. Eighty-two participants were testingIn contrast to the previous experiment, the FAST detected no relation between the test stimuli on a group level, indicating a low level of prejudice on the part of the participants – at least, insofar as the specific attitude “immigrants are cheats” relates to ethnic prejudice as a phenomenon. As such, it was concluded that *immigrant* = *cheat* was not a relation in which the sample was fluent, but that those individuals who were more fluent also felt more pressure to emit social desirable responses in the testing context.

Experiment 11 utilised image stimuli to assess relations between young females and erotic images.Experiment 11 tested twenty-seven male participants. This experiment used three FASTs deployed in serial to study the differences in FAST effect when the age of the female images was varied. This design allows the FAST to explore the limits of an equivalence class. The results of this experiment showed that participants' had extreme difficulty in completing a block in which the same response was required for Tanner 3 aged females (11.5-13yrs) and erotic images, suggesting that the FAST is particularly sensitive when test stimuli are in a relation of oppositeness or difference. When the age of the female images was increased to Tanner 4 (13-15yrs) and Tanner 5(16+ yrs) this strong opposition disappeared, with the effect vanishing on the group level and varying strongly from participant to participant. This result reflects the equivocal cultural contingencies with regard to young women's sexuality and sexualisation. More importantly, this experimental preparation demonstrates the FAST's potential in analysing the stimulus relations which are the building blocks of more complex attitudes.

The results of the FAST experiments run in the final chapter provide valuable information as to the particular strengths of the FAST procedure in an applied context. In accordance with the behavioural model of the IAT (Gavin et al., 2008) the FAST detects relations which are strongly endorsed by the socioverbal community, with which participants can be expected to be highly fluent, such as "Spiders are disgusting" and "pre-teenaged females are not sexual".

The reader must be careful to note that the FAST does not purport to detect attitudes *per se*. Rather, it detects histories of relating particular stimuli, either of framing stimuli as equivalent (spiders-disgust) or opposite or distinctive (Tanner 3 females - not sexual). Because RFT conceives attitudes as complex networks of stimulus relations, it can be used to generate *a priori* predictions about the nature of a particular relational history that would give rise to an attitude (e.g., Dixon et al, 2003). Attitudes can of course be complex and multifaceted, and this is precisely why an attitude can be conceived as a network, rather than as a single relation. Nevertheless, the FAST as a research tool is designed only to probe for a single stimulus relation that may form part of that attitude. This extreme specificity is a potential strength of the FAST in terms of its future research use, as it allows researchers to probe for and isolate specific relations hypothesized to be part of a larger attitudinal relational network.

The use of other implicit tests in this context is potentially problematic. The IAT only provides a relative measure of association strength (De Houwer, 2002, Greenwald, Nosek & Banaji, 2004). An IAT trial involves both responding towards one category while responding away from the other, it is impossible to determine which of these cues is at work in controlling the participant’s correct response rate differentials across test blocks, or the extent to which one or the other cue contributes to the total effect. In a race IAT , only the question; “Is White better than Black” is answered, without fine grained detail on whether or not the test outcome derives from a participant having a history of pro-white responding or a history of anti-black responding. By contrast, the IRAP is capable of indexing the associations between individual stimulus items. For example, in a race IRAP, a D score is produced for black-good, white-good, black-bad, and black-good attitude dimensions. The IRAP is therefore a non-relative measure of relations between stimuli. That said, its structure still requires that the attitude being studied be stated in symmetrical terms.

As such, the IAT, its variants, and the IRAP provide only indirect evidence for a particular attitude construct, by providing an index of associations that are *probably* involved in that particular attitude (De Houwer, 2002). The experimenter must select broad, symmetrical terms that could be seen to encapsulate the entire attitude construct. While this provides a useful index of whether an attitude is present for a participant, the specific stimulus relations involved in the attitude may lead to different expected behavioural outcomes, as well as different intervention strategies when we seek to change certain attitudes (for instance, reducing prejudicial behaviour).

Employing a series of FASTs in these contexts can produce fine-grained pictures of the shape of a relational network. Such an application was tested in Experiment 11, with excellent and informative results. A single FAST requires little time to complete, and so can easily be deployed in such a repeated measures design. The benefits of this type of extensive “relational network mapping” could be manifold. Research resources such as the International Affective Picture System show the benefits of large sample research with regard to the properties of certain stimuli. A similarly broad program of FAST research, while certainly ambitious, could provide a detailed picture of the mean interrelations between various verbal stimuli that constitute different hypothetical attitude constructs. In effect, the FAST could provide a snapshot of cultural attitudes on a large scale. A similar project has been attended by Anthony Greenwald and his colleagues (Project Implicit) but the utility of their research is limited by the lack of theoretical clarity with regard to “implicit attitudes” as commonly conceived by IAT researchers.

General Issues

It should be acknowledged that the concept of class strength or relational strength, as represented by the SoR Index, is not very familiar in behaviour analysis but it has been considered before in several empirical and theoretical research papers. Specifically, it has been repeatedly found using large equivalence classes trained with a linear protocol (e.g., A-B-C-D-E-F), that the nodal distance between stimuli is inversely related to the probability of the transfer of response functions. In an early study of this kind, Fields, Adams & Verhave

(1993) trained the conditional discriminations AB, BC, CD, and DE, leading to the emergence of two five-member equivalence classes.A1 and A2 were then established as discriminative stimuli for two simple responses, and as expected response functions transferred to the remaining class members. However, responses were less reliably produced by the D and E stimuli compared to the B and C stimuli. The researchers referred to this as indicative of differential *relatedness* of stimuli within large equivalence classes (see also Fields et al., 1995; Fields & Watanabe-Rose, 2008, and Reilly, Whelan, & Barnes-Holmes, 2005, for an example regarding relational frames). The important contribution this work makes to the current FAST method is that it provides a rationale for measuring these very differences in degrees of relatedness across various stimulus pairs. The FAST procedure appears to measure indirectly that degree of relatedness between stimuli, without suffering from the problem of demand characteristics or social desirability (i.e., it has high procedural implicitness). Furthermore, the FAST procedure allows the researcher to calculate a numerical score for each test outcome, which can be used as a direct single unit index of the graduated strength of relation between any two stimuli (see Bortoloti & de Rose, 2009 for the description of the attempt to do something similar using semantic differential scales).

However, there was a noticeable level of variation from participant to participant with regards to SoR indices in various experiments. On the whole, participants tended towards a positive SoR indicative of the FAST being able to detect the pre-trained relations. But at this stage of investigation, it is still difficult to impossible to infer a concrete meaning from any given SoR (i.e., are higher SoRs indicative of stronger relations?). In experiments in which laboratory trained stimuli were employed, all participants received the same training procedures, but produced very different SoR scores on the FAST. Is this an issue of imprecise measurement on the part of the FAST? Or are the varying SoR indices indicative of individuals' differences with regards to their ability to complete tasks which involve shifting contingences or that require rapid responses? The results from Experiment 4, in which participants received extensive training and then produced strongly positive SoRs, suggest that to some extent, SoRs are a function of the amount of training received. However, even within this small group, there was substantial variation, which can probably be attributed to individual differences. Resolving this question will require careful research in which participants receive robust training, but at variable levels. The design of Experiment 4 could be extended to include participants who receive a variable number of matching to sample training sessions. If SoR scores are indeed a function of the amount of training received, then each group should display mean SoR scores. The same experiment should also take measures of participants' IQ or relational ability, in order to assess the probable interaction between the two factors. Intelligence is a well-established correlate of the effectiveness of implicit tests. As IQ rises (interpreted behaviour-analytically as relational flexibility; Cassidy et al., 2011), effects on IAT tests (Greenwald et al, 2003) and IRAP tests decline (O'Toole & Barnes-Holmes, 2009)

Variability in FAST results also emerged across multiple exposures to the FAST. It would appear that the acquisition rates being measured here by the FAST blocks are themselves variable. In each of the experiments in which the same relation was tested multiple times, the relation under examination was experimentally trained. The different levels of pre-FAST training across these experiments contribute to our understanding of the behavioural variability observed in FAST effects across participants. That is, during the inconsistent test block, responses which ran counter to the equivalence training were reinforced –the inconsistent blocks act as a behavioural disrupter. Across multiple runs this may have been sufficient to destabilize the trained relations during the FAST itself. In effect, the very fluency of the relations was likely being affected dynamically across and within test blocks.

The foregoing effect was also likely compounded by the complex interaction of order effects, as in many of these experiments the order of presentation (i.e. consistent block first or inconsistent block first) was randomised at the onset of each *pair* of test blocks rather than being held constant throughout. Consider, for instance, a participant who received their FAST blocks in the order; consistent, inconsistent, inconsistent, consistent, consistent, and inconsistent. In the first test block pair, the participant would probably display the expected difference in learning rate. However, proceeding into the second pair, beginning with an inconsistent block, they will then received 2 inconsistent blocks in a row (separated by Baseline 2, in which the contingencies are not relevant to the test task). Rapid responding to the contingencies of an inconsistent block will already have been established to some extent, thus quickening the learning rate for this second inconsistent block. When the subsequent consistent block begins, the formerly “inconsistent” contingencies would now be in control of responding. This same reversal may happen once again when the participant completes a consistent block (thus producing fluent responding to the original contingencies) and soon begins another.

The impact of the shifting contingencies within multiple runs of the FAST, in particular on weakly established relations, such as those we had trained in the laboratory, also suggests a possible new experimental arrangement that may be of benefit in applied settings. While the contingency switching described above was problematic to analysis of the outcome of those experiments, the same processes which confounded that analysis can be leveraged the same as any other known behavioural process. Again drawing upon the notion of employing multiple FASTs in serial, consider the following thought experiment. We wish to ascertain the absolute strength of a well-rehearsed and frequently reinforced stimulus relation. A single FAST will provide a litmus test which indicates the existence of the hypothesised history, and the SoR will provide some indication of the degree of relatedness between the two stimuli. But the truest test of the degree of relatedness may be to expose the participant to a number of FAST blocks, switching between consistent and inconsistent block contingencies in an alternating pattern. When the fluency difference across the two blocks has disappeared across several runs, we may say that the contingencies of the FAST completely control responding, and that pre-experimental contingencies are having little if any interfering effect. In effect, the FAST might be used to measure the strength of a relation by examining how many blocks of testing are required for the reinforcement contingencies in operation during that testing to eliminate all interference effects from a participant's pre-experimental history. Such a test would not work provide a metric on the basis of a participant's performance on any one pair of test blocks, but rather across blocks in a meta-analysis. I would suggest the term Meta-FAST for such a procedure.

Conclusion

The aim of this thesis was to conceive, develop, and test a novel implicit testing format from the bottom up, based on established behaviour analytic principles. The resulting test format would be transparent in its underlying functional principles, and as such, simple to run and easily modifiable for individual research needs. In the introductory chapter, research was described which would inform the construction of the basic test format, which was named the Function Acquisition Speed Test (FAST).

The test was named for the core behavioural processes leveraged by the format to detect the existence of particular verbal histories on the part of a participant. The FAST requires participants to complete a number of simple discrimination blocks. Each trial presents a single stimulus, and participants are required to learn, via corrective feedback, whether to respond with a "left" key press (e.g. press "a" or press "z") or a "right" key press (e.g. press "j" or press "m"). For two of the stimuli, a "left" key press is reinforced, while for the other two, a "right" key press is reinforced. Participants are required to continue the block until their responses are fluent - both rapid and accurate. Participants' responses must occur within a short response window, and participants are required to complete ten trials correctly in a row to end the current block. The metric of interest is the speed at which participants acquire the required level of fluent responding (i.e., the number of trials it takes them to reach the fluency criterion on each test block).

A FAST contains two "test blocks", each of which uses the same four stimuli. Two stimuli are the stimuli of interest, suspected to be related in the participant's history. The other two are novel, unrelated stimuli. In the "consistent" block, the same response is reinforced (e.g., press z) for both of the test stimuli, and the other (e.g., press m) for the unrelated stimuli. These responses are consistent with the participant's learning history, and so should quickly result in stable, high rate responding by the participant. In the second "inconsistent block", the reinforced responses are inconsistent with the participant's learning history insofar as different responses (pres Z and press M) are required for each of two previously related stimuli. In effect, the juxtaposition of current and past reinforcement contingencies during the inconsistent block functions as a type of learning *disrupter*. The current test contingencies need to overcome the "behavioural inertia" of the previous contingencies in order for fluent responding to take place and any learning criteria to be reached. In simple terms, the "inconsistent" responses are more difficult for participants to acquire. We can use the differences in rate of acquisition across the two test blocks to infer the pre-existing "strength" of the relation we are investigating.

The FAST also contains a number of "baseline" blocks, in which all four stimuli are novel and unrelated. The purpose of these blocks is to measure the rate of acquisition of the required functional responses when the participant has no history of relating the stimuli. A baseline phase demonstrates the speed at which participants can acquire fluency in the FAST task without any past contingencies interfering with the acquisition. Inclusion of a baseline measure allows us to place differences in acquisition speed across in context. For a participant who completes such tasks very quickly, a small difference between inconsistent and consistent test blocks is more meaningful that the same raw difference may be for a participant who takes a longer time overall to complete a FAST task. This relationship between differences in learning speed and baseline acquisition speed is the basis of the Strength of Relation index used to quantify the “strength” of a the inferred pre-existing relation between the stimuli of interest.

The stated goals of this thesis were to create a novel implicit testing format that would be transparent in its underlying functional principles, simple to run and easily modifiable for individual research needs. Those goals have been met, insofar as the basic procedure has been proven to detect the kinds of relations an implicit test must be able to target, both in the context of laboratory trained relations and naturally occurring relations. The underlying principles were clearly articulated from the outset, and that clarity has allowed a great deal of flexibility in how the procedure was employed to meet differing experimental needs within the thesis. Further, the range of possibilities for future application is broad and exciting. The FAST shows potential for application in basic research into the conditions under which derived relations impact upon behaviour, linking Relational Frame Theory predictions to social cognitive studies of attitudes with carefully gathered laboratory data. Also, the FAST promises to be a valuable tool for precise study of the intricacies of hypothesised relational networks. The name “FAST” also invokes an important characteristic of the testing format – it is indeed fast to administer. A typical FAST consisting of a practice block, 2 baseline blocks, and two test blocks requires approximately 12-16 minutes to administer from start to finish. Very few participants require more than 20 minutes to complete the procedure. This trait opens up many possibilities for creative experimental design in future research. While there remains much to be done in the development of the FAST as an optimised tool for use in these research contexts, the current thesis has been successful in demonstrating the efficacy of the basic format and establishing a “standard” form for this new and functionally understood “implicit” test.

References

Arntzen, E. (2006). Delayed matching to sample: probability of responding in accord with equivalence as a function of different delays *The Psychological Record*, *56*, 135–167.

Arntzen, E., Braaten, L. F., Lian, T., & Eilifsen, C. (2011). Response-to-Sample Requirements in Conditional Discrimination Procedures, *2*(2), 505–522.

Arntzen, E., Grondahl, T., & Eilifsen, C. (2010). The Effects of different training structures in the establishment of conditional discriminations and subsequent performance on tests for stimulus equivalence. *The Psychological Record*, *60*, 437–462.

Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record, 44,* 91-124.

Barnes, D., & Hampson, P.J. (1993). Stimulus equivalence and connectionism: Implications for behavior analysis and cognitive science. *The Psychological Record, 43,* 617-638.

Barnes, D., & Keenan, M. (1993). A transfer of functions through derived arbitrary and non-arbitrary stimulus relations. *Journal of the Experimental Analysis of Behavior, 59,* 61-81.

Barnes, D., Lawlor, H., Smeets, P., & Roche, B. (1996). Stimulus equivalence and academic self-concept among mildly mentally handicapped and nonhandicapped 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. *Analysis of Verbal Behavior, 8,* 1-11.

Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational frame theory and Skinner's Verbal Behavior: A possible synthesis. *The Behavior Analyst*, *23*, 69-84.

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the implicit relational assessment procedure (IRAP) as a direct measure of implicit belief. *Irish Psychologist, 32,* 169-177.

Barnes, D., & Roche, B. (1997). Relational frame theory and the experimental analysis of human sexuality. *Applied and Preventive Psychology*, *6*(3), 117–135.

Barnes, D., Browne, M., Smeets, P. M., & Roche, B. (1995). A transfer of functions and a conditional transfer of functions through equivalence relations in three to six year old children. *The Psychological Record*, *45,* 405-430.

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., 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.

Barnes-holmes, D., Regan, D., Barnes-holmes, Y., Commins, S., Stewart, I., Smeets, P. M., Whelan, R., et al. (2005). Relating Derived Relations as a Model of Analogical Reasoning : Reaction Times and Event-Related Potentials, *Journal of Experimental Analysis of Behaviour, 84*(3). doi:10.1901/jeab.2005.

Barnes-Holmes, Y., Barnes-Holmes, D., McHugh, L., & Hayes, S. C. (2004). Relational frame theory: Some implications for understanding and treating human psychopathology. *International Journal of Psychology and Psychological Therapy*, *4*(2), 355–375.

Blackledge, J. T. (2003). An introduction to relational frame theory: basics and applications. *AContext for Science with a Commitment to Behavior Change*, *3*(4), 421.

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*(2), 192–212.

Bortoloti, R., & Rose, J. C. De. (2012). Equivalent Stimuli are More Strongly Related After Training With Delayed Matching Than After Simultaneous Matching: A Study Using the Implicit Relational Assessment Procedure. *The Psychological Record*, *62*, 41–54.

Brown, A. S., Gray, N. S., & Snowden, R. J. (2009). Implicit measurement of sexual preferences in child sex abusers: Role of victim type and denial. Sexual Abuse: A *Journal of Research and Treatment*, *21*, 166–180.

Carpentier, F., Smeets, P.M., & Barnes-Holmes, D. (2003). Equivalence-equivalence as a model of analogy: Further analyses. *The Psychological Record*, *53*, 349-371.

Cassidy, S., Roche, B. & Hayes, S. C. (2011). A relational frame training intervention to raise Intelligence Quotients: A pilot study. *The Psychological Record.*

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments and Computers*, *25*, 257-271.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of consulting psychology*, *24*(4), 349–54.

D'Amato, M. R., Salmon, D. P., Loukas, E., & Tomie, A. (1985.) Symmetry and transitivity of conditional relations in monkeys (Cebus apella) and pigeons (Columba livia).*Journal of the Experimental Analysis of Behavior*,44, 35-47.

Dasgupta, N., Greenwald, A. G., & Banaji, M. R. (2003). The First Ontological Challenge to the IAT: Attitude or Mere Familiarity. *Psychological Enquiry*, 1–7.

Davey, G. C. L. (1992). Characteristics of individuals with fear of spiders. *Anxiety Research, 4, 299-314.*

De Houwer, J., (2001). A structural and process analysis of the Implicit Association test. *Journal of Experimental Social Psychology*, *37*, 443-51

Houwer, J. D. (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*(2), 115-133.

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

De Houwer, J. (2006a). Using the Implicit Association Test does not rules out an impact of conscious propositional knowledge on evaluative conditioning. *Learning and motivation*, 1–12.

De Houwer, J. (2006b). What are implicit measures and why are we using them. *The handbook of implicit cognition and addiction*.

De Houwer, J, Beckers, T., & Moors, A. (2007). Novel attitudes can be faked on the Implicit Association Test. *Journal of Experimental Social Psychology*, *43*(6), 972–978.

De Houwer, Jan, Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Theoretical claims necessitate basic research: reply to Gawronski, Lebel, Peters, and Banse (2009) and Nosek and Greenwald (2009). *Psychological bulletin*, *135*(3), 377–9. doi:10.1037/a0015328

De Houwer, J. (2011) Why the Cognitive Approach in Psychology Would Profit From a Functional Approach and Vice Versa (2011) *Perspectives on Psychological Science, 6,* 202-209.

De Jong, P., Hout, M. Van Den, & Huijding, J. (2003). Dissociations between implicit and explicit attitudes toward phobic stimuli. *Cognition & Emotion*, *17*(4), 521–545.

De Jong, P. J., & Muris, P. (2002). Spider phobia interaction of disgust and perceived likelihood of involuntary physical contact. *Journal of anxiety disorders*, *16*(1), 51–65.

Devany, 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.

Dickins, D. W., Singh, K. D., Roberts, N., Burns, P., Downes, J. J., Jimmieson, P., & Bentall, R. P. (2001). An fMRI study of stimulus equivalence.*Neuroreport*, *12*(2), 405.

Dixon, M. R., Dymond, S., Rehfeldt, R. A., Roche, B., & Zlomke, K. R. (2003). Terrorism and relational frame theory. *Behavior and Social Issues*, *12*(2).

Dugdale, N., & Lowe, C.F. (2000). Testing for symmetry in the conditional discriminations of language trained chimpanzees. *Journal of the Experimental Analysis of Behavior, 73,* 5-22.

Dymond , S., & Alonso -Álvare z, B. (2010). The selective impact of Skinner's *Verbal Behavior* on empirical research: A reply to Schlinger (2008). *The Psychological Record, 60,* 355–360.

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.

Edwards, K. (1990). The interplay of affect and cognition in attitude formation and change. *Journal of Personality and Social Psychology*, *59*(2), 202–216. doi:10.1037//0022-3514.59.2.202

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology, 83*(6), 1441-1455

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 SocialPsychology,* **69,** 1013–1027.

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53*, 345-358.

Gannon, T. A., Wright, D. B., Beech, A. R., & Williams, S. (2006). Do child molesters hold distorted beliefs? What does their memory recall tell us? *Journal of Sexual Aggression, 12,* 5–18.

Gannon , T. A. (2009). Current cognitive distortion theory and research: An internalist approach to cognition .*Journal of Sexual Aggression* , 15 , 225 – 246.

Gavin, A., Roche, B., & Ruiz, M. (2008). A behaviourally modifed and functionallky transparent Implicit Association Test to measure levels of culturally appropriate sexual categorization of children among a sample of normal males and females, 1–28.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science, 2,* 181–193.

Gawronski, Bertram, Lebel, E. P., Peters, K. R., & Banse, R. (2009). Methodological issues in the validation of implicit measures: comment on De Houwer, Teige-Mocigemba, Spruyt, and Moors (2009). *Psychological bulletin*, *135*(3), 369–72. doi:10.1037/a0014820

Gray, N. S., Brown, A. S., MacCulloch, M. J., Smith, J. & Snowden, R. J. (2005). An implicit test of the associations between children and sex in pedophiles. *Journal of Abnormal Psychology, 114,* 304–8.

Greenwald, A G, McGhee, D. E., Schwartz, J. L. K., & Attitudes, M. I. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality*, *74*(6), 1464–1480.

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

Greenwald, Anthony G, & Banaji, M. R. (1995). Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes. *Psychological Review*.

Haimson, B., Wilkinson, K. M., Rosenquist, C., Ouimet, C. & McIlvane, W. J. (2009). Electrophysiological correlates of stimulus equivalence processes. *Journal of the Experimental Analysis of Behavior, 92*, 245-256. doi:10.1901/jeab.2009.92-245

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. (1989). Nonhumans Have Not yet Shown Stimulus Equivalence. *Journal of the Experimental Analysis of Behavior*, *51*(3), 385–392.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition.* New York: Plenum.

Heagle , A. I., & Rehfeldt , R. A. (2006). Teaching perspective-taking skills to typically developing children through derived relational responding.*Journal of Early and Intensive Behavior Intervention, 3*, 1–34.

Hinton, E.C., Dymond, S., von Hecker, U., Evans, C.J., 2010. Neural correlates of relational reasoning and the symbolic distance effect: Involvement of parietal cortex. Neuroscience 168, 138–148.

Hughes, S., & Barnes-holmes, D. (2011). The Dominance of Associative Theorizing in Implicit Attitude Research: Propositional and Behavioural Alternatives. *The Psychological Record, 61*, 465–496.

Huijding, J., & De Jong, P. (2007). Beyond fear and disgust: The role of (automatic) contamination-related associations in spider phobia. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(2), 200–211.

Kamphuis, J. H., De Ruiter, C., Janssen, B., & Spiering, M. (2005). Preliminary evidence for an automatic link between sex and power among men who molest children. *Journal of Interpersonal Violence, 20*, 1351–1365.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 774-788

Keenan, M., McGlinchey, A., Fairhurst, C., & Dillenberger, K. (2000). Accuracy of disclosure and contextual control in child abuse: developing procedures within the stimulus equivalence paradigm. *Behavior and Social Issues, 10*, 1-17.

Keown, K., Gannon, T. A., & Ward, T. (2008a). The effects of visual priming on information processing in child sexual offenders. *Journal of Sexual Aggression, 14,* 145–159.

Keown, K., Gannon, T. A., & Ward, T. (2008b). What were they thinking? An exploration of child sexual offenders' beliefs using a lexical decision task. Psychology, *Crime & Law*, *14*, 317–337.

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the Implicit Association Test: A diffusion-model analysis. *Journal of Personality and Social Psychology, 93*(3), 353–368.

Kohlenberg, B. S., Hayes, S. C., & Hayes, L. J. (1991). The transfer of contextual control over equivalence classes through equivalence classes: a possible model of social stereotyping. *Journal of the Experimental Analysis of Behavior*, *56*(3), 505–518.

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A.G. (2007). Understanding and using the Implicit Association Test: IV: Procedures and validity. In B. Wittenbrink, & N. Schwarz (Eds.), *Implicit measures of attitudes: Procedures and controversies (pp. 59-102).* New York: Guilford Press.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N., (1999). *International Affective Picture System: Instruction manual and affective ratings. Technical Report A-4*. The Centre for Research in Psychophysiology, University of Florida.

Laws, D. R. & Gress, C. L. Z. (2004). Seeing things differently: the viewing time alternative to penile plethysmography. *Legal and Criminological Psychology, 9,* 183–96.

Leslie, J. C., Tierney, K. J., Robinson, C. P., Keenan, M., & Et Al. (1993). Differences between clinically anxious and non-anxious participants in a stimulus equivalence training task involving threat words. *Psychological Record*, *43*, 153–161.

Lionello-DeNolf, K. M., & Urcuioli, P. J. (2002). Stimulus control topographies and tests of symmetry in pigeons. *Journal of the Experimental Analysis of Behavior,78,* 467–496.

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

Lipkens, R., Kop, P. F. M., & Matthijs, W. (1988). A test of symmetry and transitivity in the conditional discrimination performances of pigeons. *Journal of the Experimental Analysis of Behavior*, 49, 395-409.

Lowe, C. F. (1979). Determinants of human operant behaviour. In M. D. Zeiler & P. Harzem (Eds.), Advances in analysis of behaviour: Vol. 1. Reinforcement and the Organization of behaviour (pp. 159-192). Chichester, England: Wiley.

Malamuth, N. M., & Brown, L. M. (1994). Sexually aggressive men's perceptions of women's communications. *Journal of Personality and Social Psychology, 67,* 699–712

McFarland, S. G.,& Crouch Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition, 20,* 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*, 721-747.

Petty, R. E., Unnava, R. H., & Strathman, A. J. (1991). Theories of attitude Change. *Handbook of consumer behavior*, 241-280.

McIntire, K. D., Cleary, J., & Thompson, T. (1987). Conditional relations by monkeys: Reflexivity, symmetry, and transitivity. *Journal of the Experimental Analysis of Behavior*, 47, 279-285.

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.

Mihailides, S., Devilly, G. J., & Ward, T. (2004) Implicit cognitive distortions and sexual offending. *Sex Abuse, 16*, 333-350.

Monteith, M. J., Voils, C. I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial biases.*Social Cognition*, *19*(4), 395-417.

Moran, L., Stewart, I., McElwee, J., & Ming, S. (2010). Brief report: The training and assessment of relational precursors and abilities (TARPA): A preliminary analysis. *Journal of autism and developmental disorders*, *40*(9), 1149-1153.

Moxon, P., Keenan, M., & Hine, L. (1993). Gender-role stereotyping and stimulus equivalence. *The Psychological Record, 43*(3), 381–393.

Murphy, C., & Barnes -Holmes , D. (2009). Establishing derived manding for specific amounts with three children: An attempt at synthesizing Skinner's *Verbal Behavior* with relational frame theory. *The Psychological Record, 59*, 75–91.

Nosek, B. a, Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality & social psychology bulletin*, *31*(2), 166–80. doi:10.1177/0146167204271418

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social cognition*, *19*(6), 625–666.

Nunes, K. L., Firestone, P. & Baldwin, M. W. (2007). Indirect assessment of cognitions of child sexual abusers with the Implicit Association Test. *Criminal Justice and Behavior*, *34*, 454–75.

Ó Ciardha, C. & Gormley, M. (In press). Using a pictorial modified Stroop task to explore the sexual interests of sexual offenders against children. *Sexual Abuse: A Journal of Research and Treatment.*

Öhman, A., Dimberg, U., & Öst, L. G. (1985). Animal and social phobias: Biological constraints on learned fear responses. *Theoretical issues in behavior therapy*, 123-175.

O'Hora, D., Pelaez, M., & Barnes-Holmes, D. (2005). Derived relational responding and performance on verbal subtests of the WAIS-III. *The Psychological Record*, *55*, 155-175.

Ottaway, S. A., Hayden, D. C., & Oakes, M. A. (2001). Implicit attitude and racism: Effect of word familiarity and frequency on the Implicit Association Test. *Social Cognition*, *19*, 97–144.

O., Toole, C. O., Barnes-Holmes, D., Murphy, C., O' Connor, J., & Barnes-Holmes, Y. (2009). Relational Flexibility and Human Intelligence : Extending the remit of Skinners Verbal Behavior. *International Journal of Psychology and Psychological Therapy*, *9*(1), 1–17.

O'Toole, Catriona and Barnes-Holmes, Dermot (2009) *Three Chronometric Indices of Relational Responding as Predictors of Performance on a Brief Intelligence Test: The Importance of Relational Flexibility.* The Psychological Record, 59 (1). pp. 119-132.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of personality and social psychology*, *89*(3), 277.

Plaud, J. J. (1995). The generalized expectancy bias: An explanatory enigma. *The Behavioral and Brain Sciences, 18,* 311-312.

Power, P., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2010). The Implicit Relational Assessment Procedure (IRAP) as a measure of implicit relative preferences: A first study. *The Psychological Record*, *59*(4), 7.

Rachman, S. (1966). Sexual fetishism: An experimental analogue. *The Psychological Record, 16,* 293-296

Rehfeldt, R. A., Dillen, J. E., Ziomek, M. M., & Kowalchuk, R. K. (2007). Assessing Relational Learning Deficits in Perspective-Taking in Children with High-Functioning Autism Spectrum Disorder. *Psychological Record*, *57*(1), 23.

Ridgeway, 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 derived difference relation. *The Psychological Record, 46,* 451-475.

Roche, B., & Barnes, D. (1997). A transformation of respondently conditioned function in accordance with arbitrarily applicable relations. *Journal of the Experimental Analysis of Behavior, 67,* 275-301.

Roche, B., & Barnes, D. (1998). The experimental analysis of human sexual arousal: Some recent developments. *The Behavior Analyst, 21,* 37-52.

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*(2), 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 behavior. *The Behavior Analyst*, *25*, 75-91.

Roche, B., O'Riordan, M., Ruiz, M. & Hand, K. (2005). A relational frame approach to the psychological assessment of sex offenders. In M. Taylor & E. Quayle (Eds.), *Viewing Child Pornography on the Internet: Understanding the Offence, managing the Offender, and Helping the Victims* (pp. 109-125). Dorset: Russell House Publishing.

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*(1), 31–40.

Schlund, M. W., Cataldo, M. F., & Hoehn-Saric, R. (2008). Neural correlates of derived relational responding on tests of stimulus equivalence. *Behav Brain Funct*, *4*(6), 2120-2130.

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. Boston, MA: Authors Cooperative, Inc.

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

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M.E. Zeiler (Eds.), *Analysis and integration of behavioural units* (pp. 213-245).Hillsdale, NJ: Laurence Erlbaum Associates.

Sidman, M. (1994). *Equivalence relations and behavior: A research story.* Boston: Authors Cooperative.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: an expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, *37*(1), 5–22.

Skinner, B. F. (1957). *Verbal Behaviour*. Acton, MA: Copley Publishing Group.

Snowden, R. J., Craig, R. L., & Gray, N. S. (2011). Indirect behavioral measures of cognition among sexual offenders. *Journal of sex research*, *48*(2-3), 192–217. doi:10.1080/00224499.2011.557750

Snowden, R., Wichter, J. & Gray, N. (2008). Implicit and explicit measurements of sexual preference in gay and heterosexual men: a comparison of priming techniques and the implicit association task. *Archives of Sexual Behavior*, *37*, 558–65.

Stewart, I., Barnes-Holmes, D., & Roche, B. (2004). A Functional-Analytic Model of Analogy Using the Relational Evaluation Procedure. *Psychological Record*, *54*(4), 531.

Stewart, I., Barnes-Holmes, D., & Weil, T. (2009). Training analogical reasoning as relational responding. *Derived relational responding: Applications for learners with autism and other developmental disabilities. Oakland, CA: New Harbinger*.

Szymanski, J., & O'Donohue, W. (1995). Fear of spiders questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, *26*(1), 31-34.

Tanner, J. M. (1973). Growing up. *Scientific American, 229*, 34–43.

Tyndall, I. T., Roche, B., & James, J. E. (2004). The Relation between Stimulus Function and Equivalence Class Formation. *Journal of the Experimental Analysis of Behavior*, *81*(3), 257–266.

Tyndall, I. T., Roche, B., & James, J. E. (2009). The interfering effect of emotional stimulus functions on stimulus equivalence class formation : implications for the understanding and treatment of anxiety. *European Journal of Behaviour Analysis*, *10*(2), 215–234.

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

Weil, T. M., Hayes, S. C., & Capurro, P. (2011). Establishing a deictic relational repertoire in young children. *The Psychological Record*, *61*(3), 5.

Wigboldus, D. H. J., Holland, R. W., & Van Knippenberg, A. (2004). Single target implicit associations. *Unpublished manuscript*.

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.

Vaughan, W. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 36-42.

Villatte , M., Monest ès, J. L., McHugh , L., Frei xa i Baqu é, E., & Loas , G. (2010a). Adopting the perspective of another in belief attribution: Contribution of relational frame theory to the understanding of impairments in schizophrenia. *Journal of Behavior Therapy and Experimental Psychiatry, 41*, 125–134.

Villatte , M., Monest ès, J. L., McHugh , L., Frei xa i Baqu é, E., & Loas , G. (2010b). Assessing perspective taking in schizophrenia using relational frame theory. *The Psychological Record, 60*, 413–436.

Von Hippel, W., Brener, L., & Von Hippel, C. (2008). Implicit prejudice toward injecting drug users predicts intentions to change jobs among drug and alcohol nurses. *Psychological Science*, *19*(1), 7.

Yerkes, R. M., & Yerkes, D. N. (1928). Concerning memory in the chimpanzee. *Journal of Comparative Psychology (1921)*, *8*(3), 237-271.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Zajonc, R. B., & Markus, H. (1982). Affective and cognitive factors in preferences. *Journal of Consumer Research*, 123-131.

Appendix 1

Generic Consent Form (Used in all experiments except where noted below)

The study in which you are being asked to participate is being conducted by _____ under the supervision of Dr. Bryan Roche at the Department of Psychology at the National University of Ireland, Maynooth.

In agreeing to participate in this research you confirm that you understand the following:

This research is being conducted by ________. It is the responsibility of the above researcher to adhere to ethical guidelines in their dealings with participants and the collection and handling of data. If you have any concerns about participation you may refuse to participate or withdraw at any stage.

You have been informed as to the general nature of the study and are satisfied that you have been provided with enough information.

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

At the conclusion of your participation, any questions or concerns you have will be fully addressed. You may withdraw from this study at any time, and may withdraw your data at the conclusion of your participation if you still have concerns. Even if you turn up to participate in the study you are free to terminate your procedure in the study at any time.

By agreeing to participate in the study you are confirming that you are over 18 years of age.

Any concerns you may have after the study is completed will be dealt with by Dr. Bryan Roche at the Department of Psychology at the National University of Ireland, Maynooth.Tel: (01) 708 6026. E-mail: Bryan.T.Roche@nuim.ie.

Signed: ____

Date: __________________________

Consent Form – Experiment 2

The study in which you are being asked to participate is being conducted by _____ under the supervision of Dr. Bryan Roche at the Department of Psychology at the National University of Ireland, Maynooth.

In agreeing to participate in this research you confirm that you understand the following:

This research is being conducted by ________. It is the responsibility of the above researcher to adhere to ethical guidelines in their dealings with participants and the collection and handling of data. If you have any concerns about participation you may refuse to participate or withdraw at any stage.

You have been informed as to the general nature of the study and are satisfied that you have been provided with enough information.

You also understand that as a requirement of participating in this study you will be exposed to images which some people may find somewhat distasteful. These photographs depict images of people in distress, bodily fluids, dangerous animals and still images of violent assaults.

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

At the conclusion of your participation, any questions or concerns you have will be fully addressed.You may withdraw from this study at any time, and may withdraw your data at the conclusion of your participation if you still have concerns. Even if you turn up to participate in the study you are free to terminate your procedure in the study at any time.

By agreeing to participate in the study you are confirming that you are over 18 years of age.

Any concerns you may have after the study is completed will be dealt with by Dr. Bryan Roche at the Department of Psychology at the National University of Ireland, Maynooth.Tel: (01) 708 6026. E-mail: Bryan.T.Roche@nuim.ie.

Signed: ____

Date: _________________________

Consent Form – Experiment 11

The study in which you are being asked to participate is being conducted by _____ under the supervision of Dr. Bryan Roche at the Department of Psychology at the National University of Ireland, Maynooth.

The research forms part of an ongoing programme that is attempting to establish new and simple forms of computer-based psychological assessment that can be used for a very wide range of purposes to test an extensive variety of skills, knowledge and attitudes. This study is a pilot project and the data gathered from it are being used for research purposes only. Your performance in the study will be entirely confidential and you will be identified in our records only by a number. We will have no other record of your identity and nobody besides the researchers will have access to data concerning your performance in the study.

In the case of this study we are interested in how people respond to images that are sexual in nature. The images you will be exposed to as part of this study, are not pornographic and are sourced from standardized psychological databases of images designed for this purpose. The images will involve some mild nudity but this will not be gratuitous. You will also be asked to categorise some images into one of your types, at the end of the experiment, and to rate them along a scale.

The results of the study will not allow us to make psychological assessments of any one individual but may allow us to identify the ranges of responses entire groups of people make to the tasks you are about to see. Your data will be entirely anonymous, and your name will not be recorded anywhere. Your consent form will be separated from your computer data file and the categorization and ratings you provide at the end of the experiment.

The study consists of responding to a series of images that will be shown on a computer screen. The study as a whole should take around twenty minutes, depending on how fast you work at each of the tasks that will be presented to you.

You will be given full instructions by the computer before you begin and you may also ask questions of the researcher after the study, at which time much more detail can be provided about the nature of the tasks used in the experiment.

If you consent to participate in this experiment you are free to withdraw at any stage if you so wish.

At the conclusion of the study, the purpose of the study will be outlined in detail and you can ask more questions regarding the overall purpose of the study at that point.. Any concerns you may have after the study is completed will be dealt with by Dr. Bryan Roche at the Department of Psychology at the National University of Ireland, Maynooth.Tel: (01) 708 6026. E-mail: Bryan.T.Roche@nuim.ie.

By agreeing to participate in the study you are confirming that you are over 18 years of age.

Signed: ____

Date: __________________________

Aversive Image Stimuli used in Experiment 2

Description	Slide No.	Valence Mean(SD)	Arousal Mean(SD)

Roaches	1274	3.17(1.53)	5.39(2.39)

DisabledChild	3300	2.74(1.56)	4.55(2.06)

DistressedFem	6311	2.58(1.56)	4.95(2.27)

Description	Slide No.	Valence Mean(SD)	Arousal Mean(SD)

AttackDog	1525	3.09(1.72)	6.51(2.25)

ElectricChair	6020	3.41(1.98)	5.58(2.01)

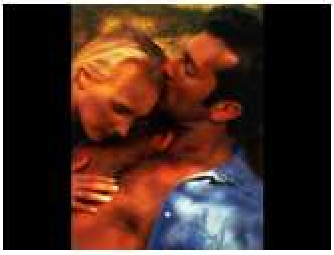
Attack	6510	2.46(1.58)	6.96(2.09)

Romantic Image Stimuli used in Experiment 2

Description	Slide No.	Valence Mean(SD)	Arousal Mean(SD)

Romance	4599	7.12(1.48)	5.69(1.94)

Romance	4606	6.55(1.62)	5.11(2.15)



Couple	4609	6.71(1.67)	5.54(2.05)

Description	Slide No.	Valence Mean(SD)	Arousal Mean(SD)

Romance	4601	6.82(1.22)	5.08(2.01

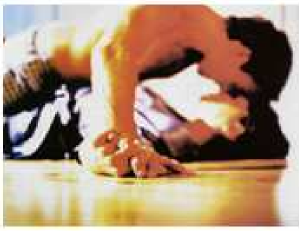
EroticCouple	4608	7.07(1.66)	6.47(1.96)

Romance	4623	7.13(1.80)	5.44(2.23)

Test Stimuli used for the FAST in Experiment 11

Erotic Stimuli (A1)

Description | Slide No. | Valence Mean(SD) | Arousal Mean(SD)



EroticCouple 4643 6.84(1.54) 6.01(2.00)

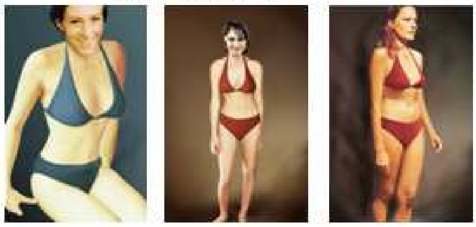
EroticCouple 4608 7.07(1.66) 6.47(1.96)

Romance 4623 7.13(1.80) 5.44(2.23)

Not Real People Stimuli (T3, T4, T5)

NRP – Tanner 3

NRP – Tanner 4



NRP – Tanner 5

Neutral Stimuli Used in FAST for Experiment 11

N1

Image	Description	Slide No.	Valence Mean(SD)	Arousal Mean(SD)
	Harvest	2515	6.09(1.54)	3.80(2.12)
	Picnic	2560	6.34(1.53)	3.49(2.07)
	Market	2597	5.61(1.26)	4.09(2.10)
	Skier	8021	6.79(1.44)	5.67(2.37)
	Skier	8030	7.33(1.76)	7.35(2.02)
	Skier	8031	6.76(1.39)	5.58(2.24)

Appendix 2:Stimuli
Nonsense Syllable Stimuli

Experiment 1

A1 ter
B1 wev
A2 hib
B2 mip
X1 ler
X2 mau
Y1 hox
Y2 yun
T1 ter
T2 wev
N1 pim
N2 kon
X3 jey
X4 huv
Y3 puh
Y4 zid

Experiment 2

A1 cug
B1 jom
N1: ter
N2: hox
N3 jey
N4 por
N5 lyr
N6 rol
X1 wev
X2 yun
Y1 vif
Y2 kon
X3 zey
X4 hib
Y3 mip
Y4 keb
X5 pim
X6 mul
Y5 arv
Y6 bix

Experiment 3
(Nx, Xx, and Yx stimuli as Experiment 2)

A1 cug
B1 luf
A2 hep
B2 paf
C1 jom
C2 mau

Experiment 4/5

A1 cug
B1 zid
A2 hep
B2 paf
C1 jom
C2 mau
P1 pim
P2 mul
Q1 arv
Q2 bix
X1 wev
X2 yun
Y1 vif
Y2 kon
X3 zey
X4 hib
Y3 mip
Y4 keb

Experiment 6/7/8

A1 cug
B1 jom
A2 zid
B2 ker
N1: ter
N2: hox
N3 jey
N4 por
N5 lyr
N6 rol

X1 wev
X2 yun
Y1 vif
Y2 kon
X3 zey
X4 hib
Y3 mip
Y4 keb
X5 pim
X6 mul
Y5 arv
Y6 bix

Experiment 7 – Practice Block Stimuli

P1 nir
P2 kyg
Q1 fam
Q2 bai
P3 cit
P4 lau
Q1 gih
Q2 lar
P5 hyf
P6 jex
Q5 gah
Q6 rec
P7 wap
P8 gif
Q7 nok
Q8 jav
P9 nud
P10 zan
Q9 mer
Q10 gyk
P11 hao
P12 fum
Q11 dur
Q12 isk
P13 sef
P14 xec
Q13 thu
Q14 pon

Appendix 3
Instructions for the stages of Equivalence Training in Experiment 4.

Stage 0 - Acquisition

In a moment some words will appear on this screen.

When the word at the top of the screen appears, you must CLICK ON IT.

Once you click on the word at the top of the screen, another word will appear at the bottom of the screen.

You must respond by clicking on the correct word at the bottom.

During this stage the computer will provide you with feedback on your performance.

You should try to get as many answers correct as possible.

If you have any questions please ask them now.

When you are ready to begin, press any key.

Stage 1

In a moment some words will appear on this screen.

When the word at the top of the screen appears, you must CLICK ON IT.

Once you click on the word at the top of the screen, TWO OTHER WORDS will appear at the bottom of the screen.

You must respond by clicking on the correct word at the bottom.

During this stage the computer will provide you with feedback on your performance.

You should try to get as many answers correct as possible.

If you have any questions please ask them now.

When you are ready to begin, press any key.

STAGE 2-3

In a moment some words will appear on this screen.

When the word at the top of the screen appears, you must CLICK ON IT.

Once you click on the word at the top of the screen, TWO OTHER WORDS will appear at the bottom of the screen.

You must respond by clicking on the correct word at the bottom.

During this stage the computer will provide you with feedback on your performance.

THIS STAGE WILL BE LONGER THAN THE LAST, AND YOU WILL REQUIRE MORE CORRECT ANSWERS TO PROCEED.

You should try to get as many answers correct as possible.

If you have any questions please ask them now.

When you are ready to begin, press any key.

STAGE 4

In a moment some words will appear on this screen.

When the word at the top of the screen appears, you must CLICK ON IT.

Once you click on the word at the top of the screen, TWO OTHER WORDS will appear at the bottom of the screen.

IN THE NEXT STAGE, THE DELAY WILL INCREASE BETWEEN CLICKING ON THE SAMPLE WORD AND THE APPEARANCE OF TWO OPTIONS AT THE BOTTOM OF THE SCREEN.

You must respond by clicking on the correct word at the bottom.

During this stage the computer will provide you with feedback on your performance.

You should try to get as many answers correct as possible.

If you have any questions please ask them now.

When you are ready to begin, press any key.

STAGE 5-6

In a moment some words will appear on this screen.

When the word at the top of the screen appears, you must CLICK ON IT.

Once you click on the word at the top of the screen, TWO OTHER WORDS will appear at the bottom of the screen.

IN THE NEXT STAGE, THE DELAY WILL INCREASE BETWEEN CLICKING ON THE SAMPLE WORD AND THE APPEARANCE OF TWO OPTIONS AT THE BOTTOM OF THE SCREEN.

THIS STAGE WILL BE LONGER THAN THE LAST, AND YOU WILL REQUIRE MORE CORRECT ANSWERS TO PROCEED.

You must respond by clicking on the correct word at the bottom.

During this stage the computer will provide you with feedback on your performance.

You should try to get as many answers correct as possible.

If you have any questions please ask them now.

When you are ready to begin, press any key.

STAGE 7

In a moment some words will appear on this screen.

When the word at the top of the screen appears, you must CLICK ON IT.

Once you click on the word at the top of the screen, TWO OTHER WORDS will appear at the bottom of the screen.

IN THE NEXT STAGE, THE DELAY WILL INCREASE BETWEEN CLICKING ON THE SAMPLE WORD AND THE APPEARANCE OF TWO OPTIONS AT THE BOTTOM OF THE SCREEN.

You must respond by clicking on the correct word at the bottom.

During this stage the computer WILL NOT PROVIDE FEEDBACK ON EVERY TRIAL.

You should try to get as many answers correct as possible.

If you have any questions please ask them now.

When you are ready to begin, press any key.

STAGE 8-9

In a moment some words will appear on this screen.

When the word at the top of the screen appears, you must CLICK ON IT.

Once you click on the word at the top of the screen, TWO OTHER WORDS will appear at the bottom of the screen.

THIS STAGE WILL BE LONGER THAN THE LAST, AND YOU WILL REQUIRE MORE CORRECT ANSWERS TO PROCEED.

You must respond by clicking on the correct word at the bottom.

During this stage the computer WILL NOT PROVIDE FEEDBACK ON EVERY TRIAL.

You should try to get as many answers correct as possible.

If you have any questions please ask them now.

When you are ready to begin, press any key.

Equivalence Testing

In a moment some words will appear on this screen.

When the word at the top of the screen appears, you must CLICK ON IT.

Once you click on the word at the top of the screen, TWO OTHER WORDS will appear at the bottom of the screen.

You must respond by clicking on the correct word at the bottom.

During this stage the computer WILL NOT provide you with feedback on your performance.

You should try to get as many answers correct as possible.

If you have any questions please ask them now.

When you are ready to begin, press any key.