Establishing and Eliminating Implicit Association Test Effects in the Laboratory: Extending the Behavior-Analytic Model of the IAT

Iseult Ridgeway[1], Bryan Roche[1], Amanda Gavin[2], and Maria R. Ruiz[3]
[1]National University of Ireland, Maynooth, [2]University of Teeside, and [3]Rollins College, Florida

In previous research, Gavin, Roche and Ruiz (2008) demonstrated that Implicit Association Test effects can be modeled in the laboratory in the form of a test that establishes competing contingencies for derived relational responding. The current study replicates and extends this finding by firstly establishing a laboratory-controlled IAT effect using nonsense syllables as stimuli, and then eliminating that effect with a simple experimental intervention. Eight subjects were exposed to an equivalence training procedure that led to the formation of two three-member equivalence relations each containing three nonsense syllables. Subjects were then exposed to a word-picture association training phase in which one member of each of the equivalence relations, printed in blue or red font, was paired with either plant or animal images, respectively. Subjects were then exposed to a modified Implicit Association Test whose outcome was successfully controlled by the organization of the trained relations. Following a stimulus equivalence re-training procedure that reorganized the stimulus equivalence relations, the IAT effect was eliminated for five of the six subjects who showed reorganization of the equivalence class and the associated derived transfer of functions. These findings lend additional support to a behavior-analytic account of the Implicit Association Test.

Key words: Implicit Association Test, derived stimulus relations, stimulus equivalence, implicit cognition

---

The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) is a social-cognitive test that is said to measure unconscious or implicit cognition while simultaneously preventing those taking the test from consciously controlling test outcomes (Banse, Seise, & Zerbes, 2001; Kim, 2003). When taking an IAT, a subject responds to a range of stimuli presented individually on a computer screen that can be classified into four categories; usually two representing a concept, such as race (e.g., black and white) and two representing an attribute (e.g., pleasant and unpleasant). Subjects are required to respond rapidly with a right-hand key press to items representing one concept and one attribute (e.g., white and pleasant), and with a left-hand key press to items from the remaining two categories (e.g., black and unpleasant). Subjects then perform a second task in which the requirements are switched (e.g., such that black and pleasant share a response and white and unpleasant share a response). The IAT compares the latencies and accuracies of responses to these two tasks. Subjects are assumed to respond more rapidly and accurately when the concept and attribute sharing the same response (e.g., left-hand key press) are pre-experimentally strongly associated (e.g., white and pleasant) than when they are weakly associated (e.g., black and pleasant).

Authors' note: The current research was conducted as an undergraduate psychology research project by the first author under the supervision of the second author. Send all correspondence regarding this article to the second author at Bryan.T.Roche@nuim.ie

Using such a testing paradigm, researchers make claims about individuals' levels of unconscious racial or other biases.

Not surprisingly, the IAT has met with some theoretical criticism, most of which concerns its mentalistic assumptions regarding core processes and the opaque nature of its scoring process (e.g., DeHouwer, 2006; Fiedler, Messner, & Bluemke, 2006; Govan, & Williams, 2004; Karpinski, & Hilton, 2001, Olson, & Fazio, 2003; Rothermund, & Wentura, 2004; Steffens & Plew, 2001; see also Gregg 2008 for a concise review). Surprisingly, however, the test has not been met with much critical empirical research compared to the overwhelming volume of research that continues to simply employ the IAT "as-is", with all of the standard test format features and data analysis techniques (e.g., Greenwald, Nosek, & Banaji, 2003) fully intact.

The IAT should be of interest to behavior analysts because it appears to function as an easily administered and subtle test for verbal relations (see Roche, O'Riordan, Ruiz & Hand, 2005), something that behavior analysts have long sought to develop. Indeed, several researchers have examined the efficacy of a behavior-analytic "implicit" test based on the concept of stimulus equivalence first proposed by Watt, Keenan, Barnes & Cairns (1991). Those researchers showed that subjects are less likely to form derived stimulus equivalence relations between two words, when those words are incompatible in the social history of the subject. More specifically, that study established that a sample of residents from Northern Ireland were often unable to derive simple equivalence relations between Catholic and Protestant symbols, whereas non-residents could do so with little difficulty. The important insight derived from this study is that the rate at which stimulus relations can be established for a given subject may function as a metric for the pre-existence of that relation in the social environment of the individual, or as a metric of its degree of compatibility with other already-established stimulus relations (e.g., between social categories). The test format is "implicit" to the extent that subjects are unable to verbally discriminate the very histories assessed using the stimulus equivalence method.

The stimulus-equivalence-based implicit test has been used by behavioral researchers to detect histories not easily ascertained through verbal reports. Examples include; cultural sectarianism (Watt et al., 1991); academic self-concept (Barnes, Lawlor, Smeets & Roche 1996; Merwin & Wilson, 2005); a history of sexual abuse (Keenan, McGlinchey, Fairhurst, & Dillenberger, 2000); a history of sex offending (Roche et al., 2005); generalized anxiety disorder (Leslie, Tierney, Robinson, Keenan, Watt, & Barnes, 1993), stereotypes regarding terrorism (Dixon, Rehfeldt, Zlomke, & Robinson, 2006) and sexual stereotyping (Kohlenberg, Hayes, & Hayes, 1991; Moxon, Keenan, & Hine, 1993; Roche & Barnes, 1996). Interestingly the stimulus-equivalence based approach to "implicit" testing preceded the advent of the IAT by several years (see Roche et al., 2005 for a detailed review of this procedure). Despite the potential of this test format for assessing verbal and social histories in a non-invasive manner, it has never been developed further or modified to appeal to a wider range of psychologists. Recent research (Gavin et al., 2008; Roche et al., 2005), however, has suggested that the core process identified by Watt et al., (1991) is in essence the same process harnessed by the IAT, albeit one not easily visible through the IAT presentation format or the associated social-cognitive nomenclature (see also O'Toole & Barnes-Holmes, 2007).

According to Roche et al. (2005) the verbal categories employed in an IAT can be conceived as equivalence classes containing words from the vernacular. In addition, higher order verbal classes also exist in the verbal repertoire of individuals that each incorporate a pair of the verbal categories employed in the IAT. For instance, for a white supremacist the verbal categories *Black* and *Bad* might participate in a higher order equivalence relation which might be referred to as "things I don't like". The IAT works by measuring the ease with which a common response function (e.g., press a left-hand key) can be established for two or more members of this higher order equivalence relation compared to members of different and unrelated equivalence relations for a given individual (e.g., *Black* and *Good*). Behavior analysts have good reason to predict such an outcome, because previous studies have shown that the existence of functional and equivalence classes between a range of stimuli interferes with the acquisition of new classes based on the reorganization of those stimuli (see Roche, Barnes & Smeets, 1997; Tyndall, Roche & James, 2004; 2009). Thus, Roche et al. suggested that the IAT operates as an indirect test for the strength of relations between verbal relations. In more technical terms, the IAT works by measuring differences in the rates of acquisition of simple functional stimulus classes under conditions in which those functional classes are either congruous (i.e., relationally consistent) or incongruous (i.e., relationally inconsistent) with socially established verbal or other stimulus relations. That is, functional response classes are established more readily when they are congruous with already-established (i.e., pre-experimental) stimulus relations. In this way, the verbal relations established in the social history of an individual can be ascertained indirectly or "implicitly" (i.e., without any form of direct questioning).

The Roche et al. (2005) behavioral model treats the concept of implicitness as referring to the fact that in the IAT format the contingencies controlling responding are not easily discriminable by the subject (i.e., they are outside conscious awareness). That is, the conflict between experimental contingencies (e.g., press left for *Black* and *Good*) and those operating in the subject's social environment (e.g., in which *Black* and *Good* are not often associated) was established across a long and complex social history. The IAT serves only to bring this conflict to bear in the experimental context. Given the extent of an adult subject's pre-experimental social history, it is very unlikely that a subject will perform in the IAT with equal fluency on trials that conflict with and do not conflict with their social history, even if they do succeed in correctly discriminating the test's core process. Consequently, the IAT allows the researcher to gather information about a subject's history without compromising test results with social desirability effects or general experimental demand characteristics. In short, rather than operating as a test for unconscious activity, the IAT appears to operate as a subtle test for functional and derived (verbal) stimulus relations operating in the social environment of the individual.

In a recent empirical test of this behavioral model, Gavin et al. (2008) exposed subjects to a word-picture association training phase in which each of two arbitrary nonsense syllables printed in blue and red font, respectively, were paired with either sexual or aversive photographic images. Subjects were then exposed to an equivalence training procedure that led to the formation of two three-member equivalence relations, each containing one of the two nonsense syllables in their respective color fonts, and two novel nonsense syllables in black font. In effect, equivalence class 1 (blue) was associated with sexual images, while equivalence class 2 (red) was associated with aversive images. Subjects were then exposed to a two-block test in which sexual and aversive images and all members of the trained equivalence relations, presented in black font, were employed. In one block, subjects were instructed to produce responses that were compatible with their laboratory history. Specifically, subjects were instructed to produce the same operant response on a computer keyboard upon the presentation of both sexual images and members of equivalence class 1 (blue), and to produce another common response upon the presentation of aversive images and members of equivalence class 2 (red). In the second block of the test the instructions were juxtaposed such that subjects were required to produce common responses to members of classes that were not previously associated with one another (e.g., sexual images and members of equivalence class 2, red). Differences in the fluency of performances across both blocks of the final test were sensitive to subjects' relational and conditioning histories. That is, subjects produced significantly more correct responses during block 1 of the test compared to block 2. Thus, Gavin et al. produced the first fully laboratory-controlled analog of the IAT using only the concepts of respondent and operant conditioning and derived stimulus relations, and without recourse to the usual mentalistic explanatory concepts (i.e., "attitudes", "mental associations", "unconscious bias").

Before we describe the current study it is important to first outline some important behavior-analytic modifications we have made to the traditional IAT (e.g., the Greenwald et al., 1998 format). Firstly, we do not present corrective feedback during the test as is traditional in the IAT. The effect, and perhaps purpose of the feedback technique is to artificially lengthen response times recorded for the inconsistent (i.e., difficult) trials *only* (i.e., in line with hypotheses) and thereby exaggerate the IAT effect. More specifically, in a traditional IAT, response variability is eliminated by forcing a correct response on each trial. This is achieved through the presentation of corrective feedback following incorrect responses only, and the instructed requirement that subjects produce the correct response before a trial can terminate. Response times are not in fact measured to the point of first incorrect response, but to the production of the forced second *correct* response. In effect, response time measures include the time taken to produce the first response, *plus* the time taken to respond privately to the on-screen feedback that an incorrect response has been made, *plus* the time taken to produce the altered response. Because errors are observed more frequently on inconsistent trials, more artificially lengthened response times are recorded for the inconsistent trial block, thereby exaggerating a reaction-time based IAT effect. Thus, response accuracy underlies the widely reported reaction-time based IAT effects, even though they are not typically reported in these terms. In line with behavioral tradition in the field of derived stimulus relations (see Whelan, 2008), the imbalanced and opaque corrective feedback procedure has been removed from the IAT for the purposes of the current research. Moreover, we emphasize response accuracy over response time measures.

In addition to employing imbalanced stimulus control across trial types, in the IAT responses times below 300ms and above 3000ms are typically truncated up to 300ms and down to 3000ms, respectively, to reduce data spread for the purpose of group-level inferential statistical analysis. However, response windows are in fact infinite on each trial and the effect of time taken beyond 3000ms to respond on individual trials (i.e., practice) has an unknown effect on the probability and speed of responses on subsequent trials. Thus, rather than contrive a narrow range of reaction times statistically, and rather than rely merely on experimental instructions to create rapid responding, our modified IAT employs a finite response window of 3000ms. In effect, we employ a type of Differential Reinforcement of High Rate (DRH) behavior technique to constrain response times and increase error rates to the point where differences in error rates (i.e., or its inverse, accuracy) are observable across test blocks. Preliminary research from our laboratory has shown that such a response window produces greater behavioral variability than the infinite response window employed in the traditional IAT.

The current study aims to replicate and extend the Gavin et al. derived relations model of the IAT by not only re-creating the laboratory controlled IAT effect, but by eliminating the effect with the use of a simple laboratory intervention to reorganize the derived stimulus equivalence relations on which the IAT effect depends. Testing the idea that laboratory controlled IAT effects can be eliminated as well as established in the first instance is an important step in the further development of the Roche at al. (2005) behavioral model of the IAT.

Method

Subjects

Eleven subjects were invited to participate in this experiment. All were acquaintances of the experimenter. Three of the 11 subjects (S5, S6, and S10) failed to pass Phase 1 and were dropped from the study. The remaining 8 subjects (4 male and 4 female) were aged 20-55 years (M= 36.5). Two were University graduates (S4, S9) while three were currently studying as undergraduates at University level (S7, S8, S11). Three further subjects were educated only to high school level (S1, S2, S3).

Apparatus

All eight phases of the experiment were presented to subjects on an Apple iBook G4®™ with a 13" monitor. Stimulus presentations were controlled using the software package Psyscope (Cohen, MacWhinney, Flatt, & Provost, 1993) which also recorded all responses. Two colored abstract shapes (5cm X 5cm ink blots, one red and one blue) and 24 photographic images (12 of animals and 12 of plants) taken from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005) were employed as stimuli. The photographic images were classified by the IAPs as producing low levels of arousal and as having moderate emotional valence (see Appendix 1). The animal images used corresponded to the slide numbers; 1440, 1441, 1540, 1590, 1610, 1640, 1670, 1675, 1710, 1740, 1810, 1812. The plant images used corresponded to the slide numbers; 5000, 5001, 5010, 5020, 5030, 5200, 5500, 5731, 5740, 5750, 5811, 5849. Six nonsense syllables were also employed as stimuli. These were; Ler, Cug, Mau, Vek, Paf, and Rog. Henceforth, these nonsense syllables will be referred to using the alphanumerics; A1, B1, C1, A2, B2, and C2, respectively. Finally, two colored amorphous "blobs" (4cm X 4cm approx) were employed as red and blue sample stimuli.

Procedure

General Experimental Sequence. Before commencing each subject was required to sign a standard consent form. Subjects took part in the experiment individually. Subjects sat comfortably at a computer desk and viewed the computer screen at a distance of approximately 70cm and at eyelevel. The current experiment consisted of eight phases that were administered during a single session (approximately 60 minutes in duration). Phases 1 through 8 were presented consecutively via the computer. Subjects were alerted that each Phase was completed by the presentation of the following statement in the centre of the computer screen in Times 24 font; "Thank You. This is the end of this phase of the experiment. Please contact the Experimenter". The experimenter then initiated the next phase manually. Figure 1 describes the sequence of the experimental phases.

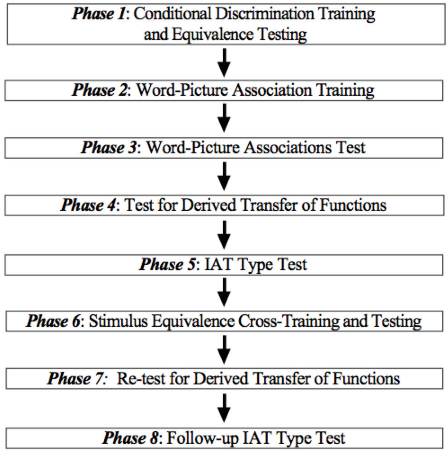

Figure 1. Schematic overview of the experimental sequence.

Phase 1. Conditional Discrimination Training and Equivalence Testing. In Phase 1, subjects were exposed to stimulus equivalence training and testing designed to lead to the formation of two-three member equivalence relations, each containing three nonsense syllables. Prior to training, subjects were presented with the following brief instructions onscreen;

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.

After reading these instructions the subject clicked on the mouse signalling the beginning of the first trial of conditional discrimination training. Tasks were presented in a quasi-random order in blocks of 16 trials (i.e., each task was presented once in a block of four which was in turn cycled four times within a block of 16, thereby preventing more than two consecutive exposures to any particular task). The sample stimuli were presented at the top center of the screen, followed 1000 ms later by two comparison stimuli presented in the bottom left and right of the screen. The right and left positions of comparison stimuli was counterbalanced across trials to prevent positional responding. All stimuli were presented in black 48-point Times font on a white background and remained on-screen until the subjects responded to one of the comparison stimuli. Responses were followed by corrective feedback delivered by the computer (i.e., the word *Correct* or *Wrong* appearing in the center of the screen in 24-point font for 2000 ms). Trials were separated by an inter-trial interval of 500ms.

The conditional discrimination phase aimed to establish four trained stimulus relations across two conditional discriminations. The training tasks were as follows: A1-B1, A2-B2, B1-C1, and B2-C2, and followed a "linear" training protocol. Subjects were exposed to successive 16-trial blocks of training until they responded correctly on 15 of the 16 tasks in a single block (i.e., 93.75% correct). In practice, however, if subjects were still engaged in training after 15 minutes the experimenter intervened, thanked the subject for their time, and excused them from further participation in the research. Immediately upon reaching criterion the subjects proceeded automatically to equivalence testing; the computer screen cleared and the following instructions were presented onscreen;

Please feel free to take a break of a few minutes at this point. When you are ready to continue, read the instructions below.
In the next phase of the experiment some more 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 or cursor. 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.
When you are ready please click the mouse button to begin the Experiment.

After clicking on the mouse subjects were exposed to a block of 16 testing trials presented in quasi-random order (i.e., four exposures to each of the four trials, with no more than two successive exposures to any task). These tasks tested for derived transitive stimulus relations; A1-C1, A2-C2, C1-A1, and C2-A2. Subjects did not receive corrective feedback during this testing period. The criterion to proceed to the next phase or to cease participation in the study was the same as in the previous phase.

Phase 2. Word-Picture Association Training. Upon reaching criterion in the stimulus equivalence testing phase, the experimenter initiated Phase 2. In this phase A1 (presented in blue) and A2 (presented in red) were associated with plant and animal photographic images, respectively. The purpose of this procedure was to establish combined color and image functions (blue/plant or red/animal) for one member of each equivalence relation. A respondent conditioning rather than a Matching-to-sample procedure was employed to analog the types of natural associative contingencies through which many words and stimuli typically acquire simple emotional and evaluative functions in the real world. Thus, the established functions might be viewed as analogous to word meanings or evaluative functions of words used in the vernacular.

At the outset, subjects were presented with the following instructions onscreen.

In a moment some words and images will appear on this screen. Your task is to look at these items carefully and to remember what you see. IT IS VERY IMPORTANT THAT YOU CONTINUE TO WATCH THE SCREEN AT ALL TIMES. 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 space bar 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 please click the mouse button.

When the subject clicked the mouse, the first trial was immediately presented. All conditioning trials were presented on the computer screen against a black background.

A trial began with the presentation of one of the two A1 or A2 nonsense syllables in 78-point Times font, appearing in the center of the screen for a period of 2000 ms. A1 was presented in blue font, whereas A2 was presented in red font. An inter-trial interval of 1000 ms followed in which the screen remained blank. After the interval a plant or animal image, randomly selected by the software program from a pool of 6 plant and 6 animal images, was presented. A1 was always followed by a plant image, whereas A2 was always followed by an animal image.

The image appeared in the center of the screen for 4000 ms, filling the screen to 95%. One second after the onset of the image, the previously displayed nonsense syllable was re-presented in the top left of the screen for the remainder of the trial (i.e., 3000 ms). While no measure was taken to ensure that subjects attended to stimuli during this phases, attention to the trials was ensured by the requirement to make an observation response at the end of each trial. That is, at the end of each trial the phrase; "Press the space bar" appeared in red in the bottom center of the screen in 24-point Times bold font and remained until the subject pressed the space bar. The space bar press functioned as an observation response that initiated the subsequent trial. The reader should also note, that if a subject failed to pay attention to stimuli during this phase they would fail the subsequent stimulus matching test (Phase 3).

There were 20 conditioning trials in total, 10 for each of the two word-picture associations (i.e., A1-plant, A2-animal). Trials were presented in a quasi-random order (i.e., one exposure to each of the two trials in a block of two and cycled 10 times), with no more than two successive exposures to any one conditional trial type. Trials were separated by a random inter-trial interval of 5, 6, 7, 8, 9, or 10 seconds, chosen randomly by the computer software. The variable inter-trial interval controlled for temporal conditioning and ensured control by the associative contingencies.

This phase established pictorial and color functions directly for A1 and A2 but relied upon the derived transfer of functions (see Dymond & Rehfeldt, 2001) for these functions to emerge spontaneously for all members of the relevant equivalence relations. This derived transfer of functions was tested in Phase 4. Finally, it is important to understand that in all subsequent phases, the A stimuli were presented only in black font.

Phase 3. Word-Picture Associations Test. Upon completion of the word-picture association phase, the experimenter manually initiated Phase 3. At the beginning of this phase subjects were presented with the following onscreen instructions;

In a moment some items will appear on this screen. Your task is to look at the item at the top of the screen and then to choose one of the items at the bottom of the screen. You should choose the item that you think is correct, by clicking on it using the mouse. If you have any questions please ask the experimenter now.

After clicking the mouse, this phase began. A trial began with the presentation of one of the two A stimuli in 48-point Times font, in the center top third of the screen. 1000ms later two comparison stimuli appeared, in the bottom left and bottom right of the screen. The position of comparison stimuli was counterbalanced across trials. All stimuli remained on the screen until the subject responded. The screen cleared immediately upon the production of a response. Trials were separated by an inter-trial interval of 1s during which the screen remained blank. Corrective feedback was not delivered by the computer during this phase. Although this phase employed a simultaneous conditional discrimination procedure, it probed for the formation of respondently conditioned stimulus relations established during the previous phase.

There were 4 trials types each presented 4 times in one mixed block of 16. The testing tasks for this phase were; A1-Blue (Red), A2-Red (Blue), A1-Plant (Animal), and A2-Animal (Plant), where the stimuli in parentheses indicate incorrect choices. Animal and plant stimuli were chosen randomly by the software program on a trial-by-trial basis (i.e., from a pool of 6 plant and 6 animal images; see Appendix 1 for the list of images used in this phase), whereas there was only one exemplar of each of the red and blue abstract shapes. Both A1 and A2 stimuli were presented in black font and all animal and plant images employed in this phase were those used during Phase 2. Subjects were required to reach a criterion of 15 out of 16 correct responses in a single block of testing to pass this phase.

Phase 4. Test for Derived Transfer of Functions. In this phase subjects were exposed to a test for derived color and image functions for the C1 and C2 stimuli. Recall that in Phase 2, A1 (presented in blue) was contingently associated with plant images, whereas A2 (presented in red) was paired with animal images. This phase tested for the transfer of both the A stimuli color and image functions to the relevant C stimuli. This phase was identical to Phase 3, except that C1 and C2 were presented in the place of A1 and A2. The four testing tasks were; C1- Blue (Red), C2- Red (Blue), C1-Plant (Animal), and C2-Animal (Plant), where the stimuli in parentheses indicate incorrect choices. Both C1 and C2 were presented in black font.

Phase 5. Baseline Implicit Association Test. In Phase 5 subjects were exposed to an IAT-type test that employed all six equivalence class members (presented in black font) and novel animal and plant images (see Appendix 1). The purpose of this phase was to see if a laboratory-controlled IAT effect could be generated with subjects who had been provided only with a history of two three-member equivalence class training and testing, and training and testing for two simple response functions.

Prior to exposure to this phase, subjects were presented with the following on-screen instructions based on those used by Greenwald et al. (1998).

In a moment some items will appear on this screen. Your task is to first look at the item and then press either the Z key on the left or the M key button on the right of the keyboard in front of you. Look now to make sure you know where they are. Use the labels at the top of the screen to help you decide which key to press.
Keep in mind:
- Keep your index fingers on the left and right buttons to enable rapid response.
- Two labels at the top will tell you which words go with each button
- Each word has a correct classification.
- Please try to go fast.
- Expect to make a few mistakes because of going fast. That's OK.
If you have any questions please ask the experimenter now.
Click any key when ready to start.

Further instructions presented on the top left and right of the screen during all tasks varied depending on the task block. The same instructions remained on-screen for the entire duration of a task block and changed only at the beginning of the second task block.

Subjects were exposed to a total of 80 trials presented across two task blocks of 40 trials (i.e., referred to here as the *consistent* and *inconsistent* test blocks). The order in which the consistent and inconsistent test blocks were presented was randomized by the computer software. Each of the two test blocks consisted of four task-types, each of which involved the presentations of one of the following stimulus types in the center of the computer screen; class 1 stimuli (A1, B1, C1), class 2 stimuli (A2, B2, C2), plant images and animal images. These four tasks were presented once each in a quasi-random order in a block of four trials and cycled 10 times (i.e., no one task was presented twice in succession).

During the consistent task block subjects were presented with two rules on the top of the screen (one on the left and the other on the right) requiring them to make a left-hand key press (i.e., press the Z key on the computer keyboard) upon the presentation of blue and plant stimuli and to make a right-hand key press (i.e., press the M key on the computer keyboard) upon the presentation of red and animal stimuli. In contrast, during the inconsistent task block, red and plant stimuli shared a left-hand key press and blue and animal shared a right-hand key press.

Both rules were presented in black 24-point Times font. A1, B1, C1, A2, B2, and C2 were presented in the center of the screen in 48-point Times black font. Sample plant and animal images were presented in the center of the screen in a 8cm x 8cm (approx) window. Figure 2 shows a sample task as it was presented to subjects.

Figure 2. Sample task presented during the IAT-type test in Phase 5. This task represents a consistent trial type.

Subjects' responses were recorded in terms of accuracy, but response times were constrained by a 3000ms response window. In effect, subjects were prevented from responding outside the 3000ms response window by the cessation of the trial and the presentation of the subsequent trial. A failure to respond within the 3000ms response window was recorded as an incorrect response. Enforcing a response fluency criterion is essential to measuring differences in response accuracy rates across the two test-blocks (see Gavin et al., for more information on this modified IAT procedure).

Phase 6. Stimulus Equivalence Cross-Training and Testing. Phase 6 was similar to Phase 1 with the difference that two of the trained relations were reversed in an attempt to eliminate the laboratory generated IAT effect. The trained relations for this phase were; A1-B2 (B1), A2-B1 (B2), B1-C1 (C2), and B2-C2 (B1), where the stimuli in parentheses indicate incorrect choices. The testing block probed for the novel derived relations; A1-C2, A2-C1, C2-A1 and C1-A2 using the same procedure as for Phase 1. It is important to understand that while the functions of the A stimuli were in no way altered, the derived functions of the B and C stimuli were expected to alter as a result of class re-orgnanization.

Phase 7. Re-test for derived transfer of functions. This phase was designed to assess if a new pattern of derived color and image function had emerged following the re-arrangement of the stimulus equivalence classes in Phase 6. It was expected that C1 and C2 should now have "swapped" the functions derived for each in Phase 4 (i.e., due to their participation in novel derived relations with A1 and A2, respectively).

Phase 8. Follow-up Implicit Association Test. In Phase 8 subjects were exposed to a follow-up IAT-type test that was identical to that presented in Phase 5. The purpose of this phase was to assess the effect of the altered stimulus equivalence relations on the size and direction of the laboratory generated IAT effect. This test was scored and analyzed in precisely the same way as the baseline IAT (Phase 5). That is, in line with the null hypothesis, it was assumed that the same IAT test effect would be observed during the follow-up IAT as during the baseline IAT. Thus, responses that were defined as correct at baseline were again defined as correct during this phase.

Results

The total number of correct responses per block of equivalence training and testing for each subject is presented in Appendix 2. Three of the eleven subjects (S5, S6, and S10) did not pass Phase 1 and only data from the remaining 8 subjects will be discussed here (but see Appendix 2). Six subjects passed the training in four or less blocks (i.e., S1, S2, S7, S8, S9 and S11). The remaining subjects, S3 and S4, required 9 and 7 blocks, respectively. Six of the 8 subjects completed equivalence testing within two blocks.

The remaining two subjects, S1 and S4, required 6 and 9 blocks of training, respectively.

All subjects who passed Phase 1 were exposed to Phase 2 during which only observational responses were required. Seven of the 8 subjects passed the test for derived color and image functions on the first block. Subject 2 required a second exposure following a score of 0/16. All eight subjects showed the predicted difference in response accuracy across test blocks. More specifically, subjects S1, S2, S3, S4, S7, S8, S9 and S11 all produced more correct responses during the consistent task block compared to the inconsistent task block, indicating that an IAT effect was modeled for all subjects.

Raw difference scores were calculated as an index of the magnitude of the IAT effect. That is, the total number of correct responses recorded during the inconsistent test block was subtracted from the total number of correct responses recorded during the consistent test block. Seven of the subjects had a difference score of +10 or greater (S1, S2, S3, S4, S7, S9 and S11). The difference score for subject 8 was +3.

A Wilcoxin Signed Rank Test showed that the difference in response accuracy across consistent (Md = 37) and inconsistent (Md = 23) blocks was significant ($z = -2.52$, $p < 0.05$). The magnitude of the difference across test blocks in terms of the variance in scores within the test blocks (i.e., the eta squared statistic) was 0.63 indicating a moderate effect using Cohen's (1988) guidelines.

All eight subjects successfully completed the equivalence cross-training and testing in Phase 6. The total number of correct responses per block of equivalence re-training and testing for each subject is presented in Appendix 3. Seven subjects passed the re-training in five or less blocks. Subject 2, however, required six blocks of training plus one further block of training following 6 blocks of unsuccessful equivalence testing. Six of the 8 subjects completed equivalence testing within two blocks. Subject 3 required five blocks of testing. Subject 2 required seven blocks of testing (interrupted by one further block of re-training).

Six of the 8 subjects (S1, S3, S7, S8, S9, and S11) responded to criterion during the test for novel derived transfer of functions in Phase 7. The two remaining subjects (S2 and S4) both produced 1/16 correct responses, indicating a complete failure to demonstrate novel derived transfer of functions patterns following equivalence re-training.

It was predicted that subjects who showed both altered equivalence class formation (Phase 6) and associated changes in the derived transfer of functions (Phase 7) should demonstrate a very much reduced or eliminated IAT effect (i.e., close to equal accuracy on consistent and inconsistent trials, or more correct responses on inconsistent trials than on consistent trials) in Phase 8.

In total, 6 of the 8 subjects (S1, S3, S7, S8, S9, S11) demonstrated the acquisition of altered derived equivalence responding (Phase 6) and the associated alterations in the derived transfer of functions (Phase 7). Only these subjects were expected to demonstrate the absence of the previously established IAT effect in the follow-up IAT test in Phase 8. Three of these 6 subjects (S1, S3, and S8) showed an IAT effect that was the reverse of that observed in Phase 5. That is, they produced more correct responses during the inconsistent task block than during the consistent block. Two further subjects (S7 and S9) produced precisely equal numbers of correct responses during both the consistent and inconsistent blocks, thereby demonstrating an eliminated IAT effect. The remaining subject (S11) continued to produce more correct responses on the consistent block compared to the inconsistent block, but this subject's difference score was reduced from 17 to 2 from the baseline to the follow-up IAT. Each subject's response accuracy during the baseline and follow-up IAT can be seen in Figure 3. The reader is reminded that during the follow-up IAT, correct responses were defined as per the baseline IAT. That is, in line with the null hypothesis, it was assumed that Phase 1-5 contingencies would control both IAT performances and that the intervention would be ineffective.

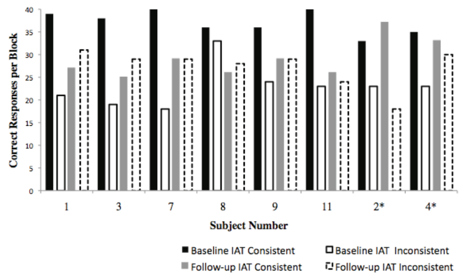

Figure 3. Response accuracies on consistent and inconsistent test blocks in the baseline and follow-up IAT-type tests (Phases 5 and 8, respectively). For both IATs the term consistent refers to a performance consistent with Phase 1-5 (i.e., baseline) contingencies. Similarly, correct responses are defined in the same way for both the baseline and follow-up IAT. * Indicates subjects who did not pass the test for new derived color and image functions in Phase 7.

Subjects 2 and 4 failed to demonstrate the acquisition of altered derived equivalence responding (Phase 6) and the associated alterations in the derived transfer of functions (Phase 7). As expected, these subjects produced more correct responses on the consistent than the inconsistent trials as they had done during the baseline IAT in Phase 5 (i.e., IAT performances were unchanged and still in line with Phase 1-5 contingencies).

As expected, a Wilcoxin Signed Rank Test revealed no statistically significant difference in response accuracy across the consistent (Md = 28) and inconsistent (Md = 29) blocks in the follow-up IAT test ($z = .00$, $p = 1$) at group level.

These data suggest that a laboratory controlled IAT effect was successfully generated using only two stimulus equivalence classes and four arbitrary stimulus functions. More importantly, experimental control was exerted over this effect, once established, using a simple intervention designed to alter the stimulus relations and derived transfer of functions upon which the laboratory-controlled IAT-effect depended.

Discussion

The purpose of the current study was firstly, to model an IAT effect in the laboratory using arbitrary stimuli and the concept of derived stimulus relations, and secondly, to assess if the effect could be disrupted with a simple intervention that altered the verbal relations on which the IAT effect depends. The study clearly replicates the findings of Gavin et al. (2008) in creating a laboratory controlled IAT effect for all 8 subjects who were exposed to the IAT. This effect was also statistically significant at the group level. The current findings also support the Roche et al. (2005) behavior-analytic position that the IAT is a measure of a subject's relative fluency in forming functional response classes that are congruent and incongruent with previously established verbal relations (e.g., stimulus equivalence classes).

The current study succeeded in completely eliminating the controlled IAT effect for five of the six subjects for whom the previously established equivalence relations and derived transfer of functions was reversed. Furthermore, the very much reduced follow-up IAT effect observed for the sixth subject (who showed evidence of altered derived stimulus equivalence relations in Phase 6), was likely also controlled by the change in stimulus equivalence contingencies implemented in Phase 6. While we might expect to see a reduction in the IAT effect size in the order of 50-80% across test exposures (see Nosek, Greenwald & Banaji, 2007), a 100% decrease in test block fluency differences is unprecedented in the research literature. Given this observation, it is highly unlikely that the current effects were due to practice across the Phase 5 and Phase 8 IATs.

The procedure used in Phase 6 of the current study was only one of two methods by which we could have undermined the laboratory-controlled baseline IAT- effect. More specifically, in order to alter the IAT performance observed at baseline, either the relation between color and image functions and the A stimuli could have been altered, or the specific equivalence class member stimuli that display these same derived functions could have been changed. The first procedure would involve targeting the respondently conditioned (non-verbal) functions of the A stimuli. For example, this could have been achieved by simply administering a further stimulus function training phase in which novel stimulus associations would be established that are incompatible with the original ones. However, the current study emerged from a primary interest in the derived relations model of the IAT proposed by Roche et al. (2005) and tested by Gavin et al. (2008). Accordingly, an intervention that would specifically undermine derived relations supporting the IAT performance was sought. Thus, the current study opted to alter the derived relations between the A and C stimuli, and thereby the derived functions of the B and C stimuli, rather than the directly conditioned functions of the A stimuli. Of course, future research may involve examining the relative ease with which laboratory-controlled IAT effects can be undermined by using the latter approach. This could be easily achieved by administering a novel stimulus association training phase in the place of the current Phase 6, in which the previously established (Phase 2) stimulus functions are juxtaposed.

The astute reader may have noticed that overall correct response rates actually dropped from the baseline to the follow-up IAT. This is a very important observation for two reasons. Firstly, this observation points to the fact that practice effects cannot easily explain the current study outcomes. That is, if one appeals to increased fluency across the to IAT test blocks to explain the reduction in differences in response accuracy across the tests blocks, then one must also expect a rise, not a reduction, in overall response accuracy. Indeed, this is what is typically observed in test-retest analyses of the IAT using real word stimuli. Secondly, this observation points to a potentially important finding regarding the acquisition of conditional discriminations and derived relations. That is, as incongruous relations are established across successive training phases, we may not necessarily expect to see that previously established relations are completely "eliminated". They may be simply under contextual control and may resurge under future conditions. However, another more interesting possibility presents itself. Baseline trained and derived relations, may in fact remain perfectly intact and may even participate in the acquisition of incompatible relations for verbally-able subjects. More specifically, it is at least possible that performances during the novel conditional discriminations presented in Phase 6 were controlled by S- rather than S+ control. That is, from the subjects' point of view, it may have been easier to respond to the reversed relations themselves, than to the entirely novel relations trained in Phase 6. Put simply, rather than learn that B2 now goes with A1 rather than A2, it may have been that subjects learned that B1 still goes with A1 but to produce the opposite response (i.e., S- control). The same would then apply at the level of derived relations. That is, subjects may not have derived that A1 and C1 no longer participate in a derived relation but that they should produce the opposite to the formerly correct response during probes for derived relation (Phase 7). In other words, the performance in Phase 6 and in subsequent phases may have been more complex in relational terms that the performance during the baseline IAT. If this is so, we should expect to see more errors overall, even though the response fluency differential between consistent and inconsistent tasks would still be necessarily reduced or even reversed. While such an account is of course speculative at this point, it should be noted that multiple covert relational responses are currently being invoked by other behavior-analytic researchers in order to account for some complex implicit test performances (see Barnes Holmes, Barnes-Holmes, Stewart & Boles, in press).

One methodological matter that might be addressed in order to improve yield rates of subjects in future research relates to the absence of a specific contingency of reinforcement during Phase 2, in which the stimulus functions were established. While a signaled observation response was required in order to maintain attention to trials during the respondent conditioning trials, it was nevertheless possible to complete this phase without attending to individual stimuli. Of course, the subsequent stimulus matching phase (Phase 3), in which the intended stimulus functions were assessed would have identified cases in which subject inattention was an issue. Nevertheless, the respondent conditioning procedure may be less effective when used with populations with learning difficulties or where multiple stimulus functions must be established. It is important to appreciate, however, that the use of a respondent conditioning preparation in Phase 2 was designed to analog the types of associative processes at work in daily life through which many simple emotional and evaluative stimulus functions are established for words. The conditional discrimination training was intended to analog the verbal processes through which those words then come to participate in rich verbal networks that support IAT test effects.

The current findings may not come as a surprise to some social-cognitive IAT researchers who are increasingly aware of how malleable IAT effects can be given only brief experimental manipulations (e.g., Gregg, Seibt, & Banaji, 2006; Mitchell, Nosek, & Banaji, 2003). However, while such researchers have examined the effects of brief stimulus associations on the creation or reduction of IAT effects, such social-cognitive research is not surprisingly couched in mentalistic terms, and outcomes are interpreted in such terms as altered propositional knowledge or unconscious associations. The current research, in contrast, provides a fully controlled experimental manipulation of the IAT effect in purely functional terms.

The level of experimental control exerted in the current study compared to that typical produced in mainstream IAT research, may also partly explain the ease with which the baseline test effects were disrupted. More specifically, social-cognitive research into the IAT invariably employs real-words as stimuli. While test effects dependent on pre-experimental learning may of course be then observed, altering such effects even with interventions designed to reorganize verbal relations is likely to be unsuccessful. This is not because the IAT measures a stable internal trait but for the simple reason that it is close to impossible to gain laboratory control over an entire natural verbal relation clearly delineated from other verbal relations. This is not least because in natural language verbal relations are large, highly contextually controlled and rarely mutually exclusive. In contrast, the current study contrived a rather simple analog of natural language over which complete control was obtained prior to both the baseline and follow-up IATs. While the current study still constitutes proof-of-principle that IAT effects can be both generated and manipulated using standard behavioral technology, future research should attempt to replicate these effects using more complex analogs of natural language.

One interesting question that arises from the current findings relates to the extent to which manipulations of underlying derived relations can lead to long-term alteration of IAT effects. More specifically, in behavioral terms, we might say that the IAT is thought to be controlled by over-arching social contingencies rather than immediate ones. Of course, this makes intuitive sense. Over-arching contingencies are pervasive and well established, whereas current contingencies are often novel and transient. Indeed, findings from the behavior analytic literature do suggest that well-established derived relations are likely to persist in the face of change efforts (see Roche et al., 1997; Tyndall et al., 2004; 2009) or even resurge following a period of time (Wilson & Hayes, 1996). It would be interesting to know, therefore, how long the current IAT effects observed post-intervention would have lasted in the absence of further stimulus equivalence training and testing contingencies. Whether or not performances would revert to baseline (i.e., an analog for over-arching contingencies) over time, or remain consistent with the intervention contingencies (i.e., an analog for control by current contingencies) is an entirely empirical question.

A complete functional-analytic investigation of the IAT is beyond the scope of any single study. The aim of the current study was merely to build upon previous research showing that IAT effects can be modeled in the laboratory using the concept of derived relations and conditioned stimulus functions alone. The current findings support that idea and would appear to constitute another important step in building a behavior-analytical understanding of the one of the most popular emerging psychological assessment tools of our times.

References

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

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

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

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.

DeHouwer, J. (2006). What are implicit measures and why are we using them. In R.W Wiers & A.W Stacy (Eds.), *The Handbook of Implicit Cognition and Addiction* (p. 11-28). Thousand Oaks, CA: Sage Publishers.

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

Dymond, S. & Rehfeldt, R. (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst, 23,* 239-254

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the "I", the "A" and the "T": Logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology, 17,* 74-147.

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

Govan, C., & Williams, K. (2004). Changing the affective valence of the stimulus items influences the I.A.T. by redefining the category labels. *Journal of Experimental Social Psychology, 40,* 357-365.

Greenwald, A., McGhee, D., & Schwartz, J. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology,* 74, 6, 1464-1480.

Greenwald, A., Nosek, B., & Banaji, M. (2003). Understanding and using the Implicit Association Test: An improved scoring algorithm. *Journal of Personality and Social Psychology,* 85, 2, 197-216.

Gregg, A. (2008). Oracle of the unconscious or deceiver of the unwitting? *The Psychologist, 21*, 9, 762-767.

Gregg, A. P., Seibt, B., & Banaji, M. (2006). Easier done than undone : Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology, 90*, 1-20.

Karpinski, A. & Hilton, J. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology,* 81, 5, 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.

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

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,* 505-518.

Lang, P., Bradley, M. & Cuthbert, B. (2005). *International affective picture system (IAPS): Digitised photographs, instruction manual and affective ratings.* Technical Report A-6. University of Florida, Gainesville, FL.

Leslie, J. C., Tierney, K. J., Robinson, C. P., Keenan, M., Watt, A., & Barnes, D. (1993). Differences between clinically anxious and non-anxious subjects in a stimulus equivalence training task involving threat words. *The Psychological Record, 43, 153-161.*

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.

Mitchell, J., Nosek, B., & Banaji, M. (2003). Contextual variations in Implicit Evaluation. *Journal of Experimental Psychology: General, 132,* 455-469.

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

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review (pp. 265–292). In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior.* Psychology Press.

Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science, 14,* 36-39.

O'Toole, C., & Barnes-Holmes, D. (2007). A derived transfer of functions and the Implicit Association Test. *Journal of the Experimental Analysis of Behavior, 88*, 263-283.

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 stimulus functions in accordance with arbitrarily applicable relations. *Journal of the Experimental Analysis of Behavior, 67,* 275-301.

Roche, B., Barnes, D., & Smeets, P. (1997). Incongruous Stimulus Pairing and Conditional Discrimination Training: Effects on Relational Responding. *Journal of the Experimental Analysis of Behavior,* 68, 143-160.

Roche, B., Ruiz, M., O'Riordan, 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.

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

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

Tyndall, I., Roche, B., & James, J. E. (2004). The relationship between stimulus function and stimulus equivalence: a systematic investigation. *Journal of the Experimental Analysis of Behavior, 81,* 257-266.

Tyndall, I., 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 Behavior Analysis, 10,* 121-140.

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

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record, 58,* 475-482.

Wilson K., & Hayes, S. (1996). Resurgence of Derived Stimulus Relations. *Journal of the Experimental Analysis of Behavior,* 66, 267-281.

Appendix 1

Valence and arousal ratings for the IAPS stimuli employed
(standardized for a combined male and female adult population)

Phase	IAPS Slide Number	Slide Name	Valence Mean	Valence SD	Arousal Mean	Arousal SD
2 and 3	5000	Flower	7.08	1.77	2.67	1.99
2 and 3	5001	Sunflower	7.16	1.56	3.79	2.34
2 and 3	5010	Flower	7.14	1.5	3	2.25
2 and 3	5020	Flower	6.32	1.68	2.63	2.1
2 and 3	5030	Flower	6.51	1.73	2.74	2.13
2 and 3	5200	Flower	7.36	1.52	2.16	2.16
2 and 3	1440	Seal	8.19	1.53	4.61	2.54
2 and 3	1441	Polar Bears	7.97	1.28	3.94	2.38
2 and 3	1540	Cat	7.15	1.96	4.54	2.35
2 and 3	1590	Horse	7.18	1.64	4.74	2.13
2 and 3	1610	Rabbit	7.82	1.34	3.08	2.19
2 and 3	1640	Coyote	6.16	1.88	5.18	1.93
5	1670	Cow	6.81	1.76	3.05	1.91
5	1675	Buffalo	5.24	1.48	4.37	2.15
5	1710	Puppies	8.34	1.12	5.41	2.34
5	1740	Owl	6.91	1.38	4.27	2.03
5	1810	Hippo	6.52	1.49	4.45	2.11
5	1812	Elephant	6.83	1.33	3.6	2.11
5	5500	Mushroom	5.42	1.58	3	2.42
5	5731	Flowers	5.39	1.58	2.74	1.95
5	5740	Plant	5.21	1.38	2.59	1.99
5	5750	Nature	6.6	1.84	3.14	2.25
5	5811	Flowers	7.23	1.6	3.3	2.33
5	5849	Flowers	6.65	1.93	4.89	2.43

Appendix 2
Total number of correct responses during each training and testing block in Phase 1.
*Indicates the final below-criterion performance before a subejct was excused from further participation in the research

Subject	**Number of Blocks**												
	1	2	3	4	5	6	7	8	9	10	11	12	13
S1: Train	11/16	9/16	9/16	16/16									
S1: Test	9/16	8/16	8/16	7/16	14/16	16/16							
S2: Train	11/16	16/16											
S2: Test	15/16	16/16											
S3: Train	8/16	12/16	10/16	9/16	14/16	15/16	15/16	15/16	16/16				
S3: Test	16/16												
S4: Train	8/16	12/16	13/16	15/16	15/16	15/16	15/16						
S4: Test	4/16	7/16	7/16	8/16	8/16	8/16	10/16	8/16	16/16				
S5: Train	10/16	12/16	9/16	8/16	8/16	9/16	11/16	12/16	11/16	13/16*			
S6: Train	12/16	15/16	13/16	13/16	15/16	16/16							
S6: Test	0/16	0/16	0/16*										
S7: Train	11/16	16/16											
S7: Test	14/16	16/16											
S8: Train	10/16	16/16											
S8: Test	16/16												
S9: Train	10/16	16/16											
S9: Test	16/16												
S10: Train	4/16	1/16	8/16	11/16	11/16	14/16	13/16	10/16	11/16	12/16	12/16	11/16	12/16*
S11: Train	10/16	15/16	14/16	16/16									
S11: Test	16/16												

Appendix 3
Total number of correct responses for each training and testing block during Phase 6.
*Indicates the final below-criterion performance before a subejct was excused from further participation in the research

Subject	Number of Blocks											
	1	2	3	4	5	6	7	8	9	10	11	12	13
S1: Train	10/16	15/16	15/16	16/16									
S1: Test	16/16												
S2: Train1	8/16	10/16	11/16	15/16	16/16								
S2: Test1	2/16	0/16	10/16	8/16	0/16	4/16							
S2: Train2	15/16												
S2: Test2	16/16												
S3: Train	12/16	16/16											
S3: Test	0/16	0/16	0/16	14/16	16/16								
S4: Train	9/16	14/16	15/16	15/16	16/16								
S4: Test	11/16	16/16											
S7: Train	11/16	16/16											
S7: Test	15/16	16/16											
S8: Train	10/16	16/16											
S8: Test	16/16												
S9: Train	11/16	16/16											
S9: Test	16/16												
S11: Train	10/16	15/16	16/16										
S11: Test	16/16												