

A FUNCTION ACQUISITION SPEED TEST FOR EQUIVALENCE RELATIONS (FASTER)

Anthony O'Reilly and Bryan Roche
National University of Ireland, Maynooth

Amanda Gavin
Teeside University

Maria R. Ruiz
Rollins College

Aoife Ryan and Glenn Campion
National University of Ireland, Maynooth

The current study employed a stimulus equivalence paradigm to assess the ability of the recently developed Function Acquisition Speed Test (FAST) to measure the existence and strength of experimentally produced derived relations. Twenty-two participants were exposed to a One-to-Many stimulus equivalence training procedure (A1–B1, A1–C1, A2–B2, A2–C2), followed by testing for derived B1–C1 and B2–C2 relations. All participants were then exposed to a FAST procedure in which a simple common operant response was established for pairs of equivalent stimuli (e.g., B1 and C1) in one block of training. In another block of training, a common response was established for pairs of non-equivalent stimuli (e.g., B1 and C2). Trial numbers required for participants to reach mastery criteria differed across the two FAST blocks, as expected, but only for those participants who had passed the prior equivalence testing phase. This finding suggests that the FAST procedure functions as both a concurrent measure of stimulus equivalence class emergence and a functional-analytic tool that might assess unreinforced and socially sensitive stimulus relations formed in the world outside the laboratory. Large inter- and intra-participant variations in performances across FAST blocks are discussed.

Key words: implicit test, Function Acquisition Speed Test, stimulus equivalence, derived relations, attitude measurement

These data were presented at the Experimental Analysis of Behaviour Group Conference, London, April 2011.

This research was conducted as part of the last two authors' undergraduate research projects and the first author's doctoral research, under the supervision of the second author.

Correspondence concerning this article should be sent to Anthony O'Reilly, Department of Psychology, John Hume Building, National University of Ireland, Maynooth, Kildare, Ireland; E-mail: anthony.g.oreilly@gmail.com

The Function Acquisition Speed Test (FAST; O'Reilly, Roche, Ruiz, Tyndall, & Gavin, 2012) is an emerging test methodology for assessing participants' histories of relational responding and stimulus relations. The FAST methodology is based on (a) the finding that already-established stimulus relations interfere with the formation of new stimulus relations and (b) the concept of behavioral momentum (Nevin & Grace, 2000).

The FAST requires participants to complete two simple discrimination test blocks. Each block utilizes the same four stimuli: two "test" stimuli related in the participant's history (or suspected to be) and two unrelated novel stimuli. Each trial presents a single stimulus, and participants must learn via corrective feedback whether to respond with a press of the "z" or "m" key on a computer keyboard. In the "consistent" block, the same response (e.g., press "z") is reinforced for both of the test stimuli, and the other response (e.g., press "m") is reinforced for the unrelated stimuli. These responses are consistent with the participants' learning history and so quickly result in stable, high-rate responding. In the second "inconsistent" block, the reinforced responses are inconsistent with the participant's learning history insofar as different responses (press "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 behavioral inertia produced by the previous contingencies in order for responding to come under experimental control and for any learning criteria to be reached. Thus, an acquisition rate difference in trial requirement to criterion is typically observed across the consistent and inconsistent test blocks in the predicted direction (i.e., most test takers reach the response fluency criterion in fewer trials on the consistent block than on the inconsistent block). The FAST, therefore, allows the researcher to identify a history of relating any two classes of stimuli.

In the first published FAST study (O'Reilly et al., 2012), participants completed a simple stimulus matching procedure that established a relation between two nonsense syllables (A1 and B1). Participants then completed a FAST, which utilized A1 and B1 and two novel nonsense syllables as stimuli (N1 and N2). In the consistent block, pressing "z" when presented with A1 or B1 as a stimulus and "m" when presented with N1 or N2 was reinforced. In the inconsistent block, pressing "m" in the presence of A1 and N1 and "z" in the presence of B1 or N2 was reinforced. Thirteen of the 18 participants reached criterion (10 correct responses in a row) more quickly on the consistent block than on the inconsistent block. The remaining five participants showed no difference in acquisition rates across the blocks or very small differences in the unexpected direction. That study provided a simple proof of concept for the FAST methodology. The purpose of the current study is to extend this demonstration by using the FAST to detect and analyze derived relations that have not been previously reinforced. If the FAST is sensitive to such relations, it may prove to be useful in applied research settings in which researchers are interested in determining the existence of derived relations between words in the vernacular (e.g., "African American" and "bad") or experimental stimuli participating in equivalence or other derived relations.

The FAST represents an extension of the stimulus equivalence-based methodology developed by Watt, Keenan, Barnes, and Cairns (1991). That technique measures the interference effects on stimulus equivalence class formation when programmed contingencies are designed to lead to emergent relations containing socially incongruous stimuli. Several studies have utilized the Watt et al. (1991) paradigm to study socially sensitive issues, such as discrimination against Middle Eastern people (Dixon, Rehfeldt, Zlomke, & Robinson, 2006), gender identity (Kohlenberg, Hayes, & Hayes, 1991; Moxon, Keenan, & Hine, 1993; Roche & Barnes, 1996), self-esteem (Barnes, Browne, Smeets, & Roche, 1995; Merwin & Wilson, 2005), and child sexual abuse (McGlinchey, Keenan, & Dillenburger, 2000; Roche, O'Riordan, Ruiz, & Hand, 2005). These studies investigated social histories that were established outside the laboratory, and all provided promise of a behavior analytic test for histories of stimulus relations. Nevertheless, this behavioral

approach to assessing stimulus relations was never harnessed into a simple and easy-to-use test format.

The O'Reilly et al. (2012) study 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 people* 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 FAST procedure is functionally very similar to a procedure utilized by Hall, Mitchell, Graham, and Lavis (2003). Hall and colleagues used a Many-to-One training procedure to train two separate associations between geometric shapes (*square*, *triangle*, *circle*, *star*) and a colored rectangle (*red* or *green*). This led to the formation of two equivalence classes (e.g., *square–red–circle* and *star–green–triangle*) and to an untrained equivalence relation between the two shapes in each class (e.g., *square–circle*). The experiment then required participants to assign a key press (either *left* or *right*) to each pair of shapes. Unlike the FAST, a between-participants experimental design was chosen. In Stage 2 of Hall's experiments, one group (the consistent condition) was required to assign the same response to stimuli that participated in the same equivalence class ("shared an associate"), and the other group was required to assign the same response to two stimuli that were not members of the same equivalence class (the inconsistent condition). In line with predictions, participants in the consistent condition produced fewer errors than those in the inconsistent condition.

Hall and colleagues (2003) explained their finding in terms of associative learning theory. In brief, this account states that when a stimulus is presented, mental representations of the associates of that stimulus are also activated in memory. In the context of the above experiment, *star* and *triangle* were separately paired with *green*. When, during testing, *triangle* was presented, it would have activated a representation of *green*. When the reinforcement was delivered for a "left" response, an associate link between *green* and a left response would have been formed, which would in turn transfer to *star* when it was presented and the *green* representation was again activated. In Experiment 4 of Hall et al., the authors attempted to exclude a verbal explanation for the observed phenomenon. In this experiment, participants were exposed to One-to-Many training in which two shapes (*snowflake* and *triangle*) were matched with two nonsense syllables (*WUG* and *ZIF*) and also with two colors (*red* and *green*). Participants were then required to categorize the nonsense syllables and colors with either a left or right key press. As before, participants were separated into consistent and inconsistent experimental groups. However, in this experiment, an additional stage consisting of two trials was added. Each of the color stimuli were presented, along with a 5-point rating scale, labeled "WUG" at one point and "ZIF" at the other, with the middle point marked "don't know." In line with Hall's predictions, verbal evaluations of the relations trained diverged from their responses during the critical testing stage. This suggested that the mental associations established during training were being measured directly and more reliably using the experimental procedure than by verbal reports and that (according to Hall et al., 2003) the process that produced the Stage 2 performances was not verbally mediated.

Smyth, Barnes Holmes, and Barnes-Holmes (2008) challenged the Hall et al. (2003) account directly with a series of elegant experiments. They demonstrated that the divergence between participants' responses to the experimental stimuli and their verbal evaluations of those stimuli was likely due to instructional control. More specifically, they

showed that instructions alone, as well as combinations of stimulus matching and instructions could produce the same results. These experiments undermine a purely associative account and suggest a role for verbal processes in acquired equivalence effects—a position that has long been held by behavior analysts (e.g., Sidman, 1994). Indeed, the phenomenon of derived relational responding is thought to be a core process of human verbal behavior (see Hayes, Barnes-Holmes, & Roche, 2001), and as such, underpins the authors' own understanding of the FAST effect.

The use of derived equivalence relations as a laboratory analog of verbal relations of interest to social researchers (e.g., in the context of attitude research) is supported by a growing body of research that 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, McCullagh, & Keenan, 1990; Devany, Hayes, & Nelson, 1986; Hayes & Hayes, 1989; Lipkens, Kop, & Matthijs, 1988). Other research has found it difficult at best 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 behavior analysts to develop models of meaning (Bortoloti & de Rose, 2009) as well as to understand grammar and syntax from a behavioral perspective (e.g., Barnes-Holmes, Barnes-Holmes, & Cullinan, 2000; Barnes & Hampson, 1993; Hayes et al., 2001; Hayes & Hayes, 1989). 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 analogue of implicit verbal relations as conceived and assessed by social-cognitive researchers.

The current study examined 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.

Method

Participants

Twenty-four participants were recruited from the undergraduate population of the National University of Ireland, Maynooth. Of the 24 who began the study, 17 of these passed equivalence training and testing, but 1 was eliminated due to failing to complete a FAST test block in less than 100 trials (see Phase 3: The Function Acquisition Speed Test). The remaining 7 participants were employed as control participants, but 1 of these was also eliminated due to failing to complete a FAST test block in less than 100 trials. Of the remaining 22 participants whose data were analyzed, 11 were male and 11 were female. Ages ranged from 18 to 48 years ($M = 24.09$, $SD = 7.909$).

Apparatus

All phases of the experiment were presented to participants on an Apple Macbook laptop computer with a 13-in monitor (1024 × 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 28 nonsense syllables (see the appendix) randomly assigned to their roles as samples, comparisons, and FAST test stimuli. These will be referred to later using alphanumeric.

General Experimental Procedure

The experiment consisted of three phases. Phase 1 (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 2 (equivalence testing) tested for the derived relations emergent from Phase 1 (i.e., C1–B1, B1–C1, B2–C2 and C2–B2). Phase 3 (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 protocol. Each of four training trials appeared eight times in a quasi-random order, for a total of 32 trials. The relations trained were A1–B1 (B2), A2–B2 (B1), A1–C1 (C2), and A2–C2 (C1), where unreinforced choices are in parentheses. 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 a computer screen against a white background. A trial began with the presentation of the sample stimulus at the top center of the screen in black 24-point font. The two comparison stimuli appeared in the bottom left and right corners of the screen 1,000 ms later. The positions of comparison stimuli were counterbalanced across trials. The stimuli remained on screen (i.e., simultaneous matching-to-sample) until the participant emitted a response (i.e., clicking on one of the comparison stimuli). The screen then cleared, and corrective feedback (CORRECT or WRONG) was displayed in red 24-point font in the center of the screen for 1,500 ms. 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. 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 four testing blocks, they were classified as control participants and they proceeded to Phase 3 as normal.

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 an additional baseline block. In this experiment, participants were exposed to four baseline blocks, one at the beginning of the phase and one after each pair of test blocks. That is, participants were exposed to the following series: Baseline 1, Test Blocks 1, Baseline 2, Test Blocks 2, Baseline 3, Test Blocks 3, Baseline 4. After the completion of each block, the instructions page for the next block appeared, allowing participants 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 participants for their participation and instructing them to contact the experimenter.

Each block (i.e., test blocks as well as baseline blocks) utilized four stimuli. In each block of the FAST, participants were 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 continued until a participant reached a predetermined criterion (a sequence of 10 responses in which the subject produces no more than one error at any point, i.e., 9/10 correct). The number of trials a participant required to reach this criterion on each was the primary datum.

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, and Y2 as stimuli, whereas Baseline 2 employed X3, X4, Y3, and 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 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 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 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 the four nonsense syllable stimuli (i.e., X1, X2, Y1, or Y2) in the center of the screen in black 48-point font. The stimuli remained onscreen for a period of 3 s or until a response was emitted (i.e., a 3-s 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 center of the screen for 1.5 s). If no response was emitted within the 3-s 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 3-s response window. Participants were exposed to trials until a criterion of 9/10 correct in a 10-trial sequence 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.

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.

Test blocks. The test blocks utilized the A1 and B1 stimuli from the equivalence training and testing phases and two additional 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.

In summary, the test 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.

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.

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 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.

Results

Data for 22 participants was analyzed (see the 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 analyzed first below.

Table 1
Number of Blocks of Equivalence Training and Testing Blocks Required to Reach Criterion by Each Participant

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

Note. F indicates that the performance resulted in a fail classification.

Experimental Participants

Participants completed four baseline blocks during the FAST procedure. Table 2 shows the number of trials required by each experimental participant (i.e., those who passed equivalence testing) 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	39.00
12	100	63	35	100	74.50
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	26.00
Mean	44	31.13	21.38	36.31	33.20
(SD)	(34.31)	(15.66)	(19.61)	(29.22)	(19.43)

These data were analyzed using a repeated-measures analysis of variance 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 toward 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 Strength of Relation index).

FAST Run 1 (critical test blocks). Table 3 outlines the trial requirements for each participant on each block of the three FASTs administered. Eleven of the 16 participants showed a faster rate of response function acquisition on the consistent block than on the inconsistent block, as expected. The average trial requirement difference in the positive direction was 24. In contrast, the difference in the unexpected direction was smaller ($M = -14$). 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 16 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 16 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. In contrast, the mean difference in the unexpected direction was smaller ($M = -6.6$). 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 for Experimental Participants

Participant	FAST 1				FAST 2				FAST 3			
	Difference	Consistent	Inconsistent	Order	Difference	Consistent	Inconsistent	Order	Difference	Consistent	Inconsistent	Order
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)			

Note. 1 = consistent block first; 2 = inconsistent block first.

The Strength of Relation index. The Strength of Relation (SoR) index is a simple calculation that 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).

An SoR index was calculated by dividing each participant's raw difference score (inconsistent block minus consistent block) by the natural logarithm of that participant's mean baseline trial requirement. This created an index that is 0 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 is 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. Single-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 0, 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 exposure. The FAST effect during this first exposure was significant using both raw difference scores and differences in SoR indices from 0.

Table 4
Strength of Relation (SoR) Indices and Block Order for Experimental Participants

Participant	FAST 1		FAST 2		FAST 3	
	SoR index	Order	SoR index	Order	SoR index	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)	

Note. A positive index score indicates a predicted FAST effect. 1 = consistent block first; 2 = inconsistent block first.

Control Participants

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)

Note. The dot indicates data lost due to computer error.

FAST Run 1 (critical blocks). 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, P18, P19, and P22), 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 4 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, three participants showed a faster rate of acquisition on the consistent block, while three others showed the opposite effect. P20 opted to end his participation after two runs of the FAST, 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 for Control Participants

Participant	FAST 1				FAST 2				FAST 3			
	Difference	Consistent	Inconsistent	Order	Difference	Consistent	Inconsistent	Order	Difference	Consistent	Inconsistent	Order
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)			

Note. 1 = consistent block first; 2 = inconsistent block first. Dots indicate data lost due to computer error.

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 0. In each instance, the SoR scores were not significantly different from 0 (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 (SoR) Indices and Block Order 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)	

Note. 1 = consistent block first; 2 = inconsistent block first. Dots indicate data lost due to computer error.

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 on 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 (as reported in O'Reilly et al., 2012). Traditionally, similar test methodologies have emerged from research in which real words taken from the vernacular are used as stimuli. However, by beginning with stimulus relations entirely under experimental control, we have been able to avoid potential confounds relating to real-word stimulus choice, word frequency, word length, and so on. Thus, the FAST can be used with some degree of confidence in future research in applied settings. More important, however, the FAST procedure was shown to be sensitive to derived relations that had never been reinforced in the history of the participants. Derived relational responding is thought to be a fundamental aspect of human verbal behavior (see Hayes et al., 2001), and as such, the current findings bode well for the FAST's future extension into testing for complex verbal and social histories.

In the current study, 11 of 16 participants who passed the equivalence test showed effects in the predicted direction on their first exposure to the FAST. Those effects that 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 toward 0, with considerable variation around that point. In effect, control participant data showed no convincing tendency toward either positive or negative FAST effects, suggesting that FAST performances were not under the control of a history of responding to the experimental stimuli.

The effects observed here for the FAST methodology are interesting for the most part only at the group level. One important issue that has to be borne in mind, however, in any critique of the data trends, is that the stimulus relations under analysis here were derived by participants from a brief stimulus equivalence training phase and were not reinforced at any time. It is to be expected, therefore, that FAST effects under these circumstances should be weaker than those observed for directly trained relations or well-rehearsed verbal relations taken from the vernacular (e.g., *evil–bad*). Viewed from this perspective, the current effects may be all the more promising insofar as the FAST procedure appears to have identified an untrained, implied (i.e., derived) relation between two stimuli (for similar observations in relation to a behavior-analytically modified IAT, see Gavin, Roche, & Ruiz, 2008; Ridgeway, Roche, Gavin, & Ruiz, 2010).

The foregoing issue notwithstanding, it is still important to confront the matter of large variances in FAST effects across participants and across runs. While group trends were always 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. Much research remains to be conducted to elucidate the sources of this behavioral variability. However, 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. One potentially large source of test performance variability, therefore, may be a lack of familiarity with the task format. Future research in our laboratories will be investigating the stabilizing influence of practice blocks upon future FAST performance.

Of the 22 participants who underwent equivalence training, 8 failed to pass the equivalence testing phase. 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. However, the current experiment did in fact employ an effective training protocol. More specifically, we 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. First, the current 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). Second, the current experiment presented the comparison stimuli 1,000 ms after the sample in every trial. However, increasing the delay between sample and comparison presentation across trials has been shown to increase the yield in equivalence responding (Arntzen, 2006). Third, the introduction of a fading of consequences across training could also be implemented, in order to control for extinction (see Arntzen et al., 2010). These potential shortcomings in the equivalence training procedure may indeed have led to a weakening of the observed FAST effects, but they cannot explain them.

The idea that the stimulus equivalence relations established in the current study were suboptimally trained may in fact contribute to our understanding of the behavioral variability observed in FAST effects across participants. That is, during the inconsistent test block, responses that 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. Therefore, future research would benefit from the use of far more robust relation training procedures involving overtraining, the fading of consequences, observation response requirements, and delayed matching-to-sample methods.

Another possible source of FAST effect variability across participants may be related to what might be viewed as a somewhat crude response recording system. More specifically, during a FAST, an error response may be recorded in one of two ways: as an incorrect response or as a missed response (i.e., because no key was pressed within the response window). Both of these forms of response were treated as equivalent in the current study. These two responses may, however, be quite functionally distinct (i.e., they reflect different forms of stimulus control) and are in fact consequted differentially. That is, in the former case, corrective feedback is delivered, whereas in the latter case, no feedback is provided at all (i.e., the failure to press a key is neither punished nor reinforced). It might be suggested, therefore, that some of the variability in FAST effects observed in the current study could be attributed to an insufficient response classification system. For instance, participants with large numbers of errors may have simply been observing multiple trials without making a response, unaware that their non-responding

was being recorded as an error on each trial. This may have interfered with learning rates on each test block. However, a post-hoc analysis of the data set showed that missed responses comprised only 2.015% of the total of all overt responses made (both correct and incorrect). In fact, no more than two non-responses were recorded on any given block for any given participant. The remaining participants made no or only one missed response per block. Further analysis showed that missed responses were the basis of only 7.26% of all recorded errors for experimental participants. Thus, clusters of missed responses do not in fact account for the error rate differences recorded across the consistent and inconsistent blocks in the current study.

Interestingly, just as more errors were recorded during inconsistent blocks than during consistent blocks (as expected), so too were more *missed responses* observed during inconsistent than during consistent blocks, as we might also expect. More specifically, analysis showed that missed responses were the basis of only 8.4% of all recorded errors during inconsistent test blocks and 5.3% during consistent test blocks. Thus, missed response rates cannot easily account for the ranges of raw response accuracy differentials observed for most participants. Nevertheless, this is an issue that itself requires empirical analysis. For instance, to develop a sound behavioral analysis of the functional nature of test responses, we need to understand the effects of controlled feedback omission and/or response omission on subsequent response patterns and ultimately on test outcomes (see Gavin, Roche, Ruiz, Hogan, & O'Reilly, 2012, for a discussion).

The order of the multiple FAST blocks may also have influenced effect directions and magnitudes in some cases and may also go some way toward explaining the apparent instability of the FAST effect across exposures. To understand how this order effect may have worked in the current study, let us reconsider the analogy of behavioral momentum. When a participant emitting stable behavior is exposed to a change in reinforcement contingencies, the new contingencies will need to overcome the "inertia" produced by the previous learning history. In the first run of the FAST here, order did not appear to be problematic (see also O'Reilly et al., 2012). However, participants were exposed to three consecutive runs of the FAST. In each run, the block order was randomized. This resulted in some participants being exposed to two consistent or two inconsistent blocks consecutively (albeit separated by a baseline block). In such a scenario, the eventual introduction of an altered reinforcement contingency would be met with greater-than-usual inertia or resistance to change (i.e., two rather than just one block of responding has taken place under a different reinforcement contingency). Consider, for instance, the performances of P9 and P10. Both of these participants showed a reversal of the FAST effect from Run 1 to Run 2. In the first run, P9 showed a difference score of +63. The block order employed in that run was consistent followed by inconsistent. On the second run of the FAST, however, the order was reversed (i.e., in line with the random block order protocol). This, in effect, resulted in P9 being exposed to the inconsistent block reinforcement contingencies twice in succession (albeit with a baseline block in between). This may partly explain why a reversal of effect was observed for this and other participants (e.g., P10) to whom the same analysis applies.

Given the foregoing, it is important that implicit test studies consider the possibly confounding effects of randomizing the order of blocks across multiple runs. However, an interesting conceptual issue arises from the foregoing analysis regarding block order within individual runs. That is, from the outset, popular tests such as the IAT have reported order effects of the same type expected when one adopts a behavioral momentum perspective, as we have done here. That is, learning under the control of contingencies compatible with those in operation in the participant's history will encounter no resistance to change. However, when contingencies subsequently change and are counter to those in operation in the participant's history, we might expect to see a relative reduction in rate of response acquisition to preset fluency criteria. In essence, this is the core perspective of a behavioral model of the IAT (see Gavin et al., 2008).

Importantly, however, in a behavioral momentum analysis, the order of events is always conceived as moving from consistent to inconsistent test blocks. Of course, it is possible to administer blocks in the opposite order, and differences in learning rates will still likely be observed. However, in this case, the difference in learning rates may be attenuated by the effects of the contingency shift itself on acquisition rates. More specifically, behavioral momentum will continue to gather across successive trials on the first (inconsistent) block. Thus, control by novel reinforcement contingencies has already been established to some extent by the time the second (consistent) block is administered. By definition, at least some resistance to change will therefore be encountered in the opposite direction (i.e., consistent with the participant's history). Now the otherwise unimpeded learning during the consistent block is in fact impeded by a brief and recent history of responding under other contingencies. This will at least reduce effect sizes and, in some cases, may even eliminate or reverse them (see Klauer & Mierke, 2005).

Interestingly, the foregoing order effect is precisely what one commonly observes in the IAT (see Lane, Banaji, Nosek, & Greenwald, 2007), although the explanations for this effect are cognitive rather than behavioral in nature (e.g., the task set switching account proposed by Klauer & Mierke, 2005). IAT researchers have tried to deal with this order effect problem by providing extensive practice involving task switching before the administration of the second block of implicit testing (see Nosek, Greenwald, & Banaji, 2005). From a behavioral perspective, such an intervention produces generalized sensitivity to shifting contingencies and will indeed have the desired effect. That is, it will reduce the effect of local behavioral momentum (e.g., responding in particular spatial locations for particular stimuli) so that only the effects of extended histories of behavioral momentum involving experimental stimuli will influence correct response rates during test blocks (IAT) or learning blocks (FAST). Future research should involve an analysis of the effects of prior generic contingency shift training (e.g., see Dymond, Cella, Cooper, & Turnbull, 2010) on the stability of FAST effects and any order effects. This would be a preferable option to merely randomizing block order as a means of accepting poor levels of experimental control. It may emerge, for instance, that there is an ideal and well-understood block order that should be used for all participants. This might seem strange to psychologists who are used to averaging out psychological effects through the use of randomization procedures wherever possible, but it would make perfect sense if the focus of research is on behavioral control and well-elucidated behavioral processes.

While there is considerable variability in FAST effects across successive runs, it is important to point out that significant or near-significant effects remain across successive exposures. Effects did tend to decrease on the whole and stabilize (i.e., in terms of standard deviation) from one exposure to the next. Nevertheless, the effect sizes were still sufficiently large and stable across participants that inferential analyses found them to be significant (Run 3) or tending toward significance (Run 2). We can conclude, therefore, that the FAST is vulnerable to practice effects across successive exposures, but not to an extent sufficient to eliminate all effects within three exposures.

Conclusion

The current experiment demonstrated that the FAST procedure is capable of detecting laboratory induced implied (i.e., derived) relations between arbitrary stimuli. Taken alongside the findings of O'Reilly et al. (2012), the evidence supporting the utility of the FAST methodology for the assessment of the existence and strength of stimulus relations is growing. Perhaps more interestingly, the FAST method was sensitive here to precisely the types of implied verbal relations of interest to social-cognitive psychologists. These findings thereby provide empirical support satisfactory to the experimental analysis of behavior that untrained relations can indeed be measured using implicit test methodology. As such, the current findings are of importance not only to behavior analysts interested in building

implicit tests (e.g., the IRAP) but to all researchers working within a social–cognitive paradigm. Much research remains to be done to hone the current method and eliminate sources of variability. Another important next step for researchers will be to test the utility of the FAST in “real world” contexts, examining relations between real-word stimuli whose relations have been established (either directly or indirectly) by the verbal community.

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*(1), 135–167.
- ARNTZEN, E., BRAATEN, L. F., LIAN, T., & EILIFSEN, C. (2011). Response-to-sample requirements in conditional discrimination procedures. *European Journal of Behaviour Analysis*, *12*, 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., 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, D., & HAMPSON, P. J. (1993). Stimulus equivalence and connectionism: Implications for behavior analysis and cognitive science. *The Psychological Record*, *43*, 617–638.
- BARNES, D., MCCULLAGH, P. D., & KEENAN, M. (1990). Equivalence class formation in non-hearing impaired children and hearing impaired children. *The Analysis of Verbal Behavior*, *8*, 19–30.
- BARNES-HOLMES, D., BARNES-HOLMES, Y., & CULLINAN, V. (2000). Relational frame theory and Skinner's *Verbal Behavior*: A possible synthesis. *The Behavior Analyst*, *23*, 69–84.
- BARNES-HOLMES, D., STAUNTON, C., WHELAN, R., BARNES-HOLMES, Y., COMMINS, S., WALSH, D., ... DYMOND, S. (2005). Derived stimulus relations, semantic priming, and event-related potentials: Testing a behavioral theory of semantic networks. *Journal of the Experimental Analysis of Behavior*, *84*, 417–433. doi:10.1901/jeab.2005.78-04
- BORTOLOTTI, R., & DE ROSE, J. C. (2009). Assessment of the relatedness of equivalent stimuli through a semantic differential. *The Psychological Record*, *59*, 563–590.
- BORTOLOTTI, R., & DE ROSE, J. C. (2012). Equivalent stimuli are more strongly related after training with delayed than with simultaneous matching: A study using the Implicit Relational Assessment Procedure (IRAP). *The Psychological Record*, *62*, 41–54.
- 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. doi:10.3758/BF03204507
- 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. doi:10.1901/jeab.1986.46-243
- 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), 405–411. doi:10.1097/00001756-200102120-00043

- DIXON, M. 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.
- 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. doi:10.1901/jeab.2000.73-5
- DYMOND, S., CELLA, M., COOPER, A., & TURNBULL, O. H. (2010). The contingency-shifting variant Iowa gambling task: An investigation with young adults. *Journal of Clinical and Experimental Neuropsychology*, *32*, 239–248. doi:10.1080/13803390902971115
- GAVIN, A., ROCHE, B., & RUIZ, M. R. (2008). Competing contingencies over derived relational responding: A behavioral model of the Implicit Association Test. *The Psychological Record*, *58*, 427–441.
- GAVIN, A., ROCHE, B., RUIZ, M. R., HOGAN, M., & O'REILLY, A. (2012). A behavior-analytically modified Implicit Association Test for measuring the sexual categorization of children. *The Psychological Record*, *62*, 55–68.
- HAIMSON, B., WILKINSON, K. M., ROSENQUIST, C., OUMET, 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*, 266–276. doi:10.1037/0096-3445.132.2.266
- HAYES, D., BARNES-HOLMES, D., & ROCHE, B. (EDS.). (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York, NY: Plenum.
- HAYES, S. C., & HAYES, L. J. (1989). The verbal action of the listener as the basis for rule-governance. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 153–190). New York, NY: Plenum Press.
- KLAUER, K. C., & MIERKE, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the IAT effect. *Personality and Social Psychology Bulletin*, *31*, 208–217. doi:10.1177/0146167204271416
- KOHLBERG, B. K., 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. doi:10.1901/jeab.1991.56-505
- LANE, K. A., BANAJI, M. R., NOSEK, B. A., & GREENWALD, A. G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far). In B. Wittenbrink & N. S. Schwarz (Eds.), *Implicit measures of attitudes: Procedures and controversies* (pp. 59–102). New York, NY: Guilford Press.
- 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*, 467–495. doi:10.1901/jeab.2002.78-467
- LIPKENS, R., KOP, P. F. M., & MATTHIJS, W. (1988). A test of symmetry and transitivity in the conditional discrimination performance of pigeons. *Journal of the Experimental Analysis of Behavior*, *49*, 395–409. doi:10.1901/jeab.1988.49-395
- 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.

- 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.
- MOXON, P. D., KEENAN, M., & HINE, L. (1993). Gender role stereotyping and stimulus equivalence. *The Psychological Record, 43*, 381–393.
- NEVIN, J. A., & GRACE, R. C. (2000). Behavioral momentum and the law of effect. *Behavioral and Brain Sciences, 23*, 73–130. doi:10.1017/S0140525X00002405.
- NOSEK, B. A., GREENWALD, A. G., & BANAJI, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin, 31*, 166–180. doi:10.1177/0146167204271418
- O'REILLY, A., ROCHE, B., RUIZ, M., TYNDALL, I., & GAVIN, A. (2012). The Function Acquisition Speed Test (FAST): A behavior-analytic implicit test for assessing stimulus relations. *The Psychological Record, 62*, 507–528.
- 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., 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, UK: Russell House Publishing.
- SIDMAN, M. (1994). *Equivalence relations and behavior: A research story*. Boston, MA: Authors Cooperative.
- SMYTH, S., BARNES-HOLMES, D., & BARNES-HOLMES, Y. (2008). Acquired equivalence in human discrimination learning: The role of propositional knowledge. *Journal of Experimental Psychology: Animal Behaviour Processes, 34*, 167–177. doi:10.1037/0097-7403.34.1.167
- WATT, A., KEENAN, M., BARNES, D., & CAIRNS, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41*, 33–50.

Appendix

Nonsense syllables used (equivalence training and testing and FAST procedure)

A1: cug	X1: wev	X3: zey	X5: pim	X7: tuk	N1: ter
A2: mau	X2: yun	X4: hib	X6: mul	X8: dit	N2: nox
B1: vek	Y1: vif	Y3: mip	Y5: arv	Y7: rit	N3: jey
B2: zid	Y2: kon	Y4: keb	Y6: bix	Y8: bam	N4: por
C1: jom					N5: lyr
C2: ler					N6: rol