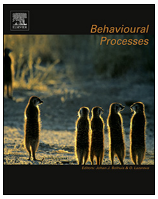
Behavioural Processes 178 (2020) 104179

Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc



Measuring differential nodal distance using the function acquisition speed test

Jamie Cummins[a,*], Bryan Roche[b]

[a] Ghent University, Department of Experimental- Clinical and Health Psychology, Belgium
[b] Maynooth University, Department of Psychology, Ireland

ARTICLE INFO

Keywords:
Function acquisition speed test
Stimulus relatedness
Nodal distance
Stimulus equivalence
Implicit measures
Derived relational responding

ABSTRACT

Increasing evidence suggests that the relatedness of stimuli within the Function Acquisition Speed Test (FAST) methodology is sensitive to the learning histories of participants. For example, this method is sensitive to differences in the amount of baseline training provided to establish stimulus equivalence relations using arbitrary stimuli (Cummins et al., 2018a). However, it has not yet been investigated whether the relatedness of stimuli within the FAST varies based on differential nodal distances between stimuli *within* stimulus classes. If so, the FAST could serve an important adjunct assessment procedure for researchers who wish not only to assess the formation of stimulus classes using traditional methods, such as matching-to-sample, but also the relative relatedness of stimuli within complex stimulus classes (i.e., nodal distance). The current study sought to investigate this possibility. Participants (n = 16) were trained in the formation of two 4-member equivalence classes consisting of arbitrary nonsense syllables. Following this, participants completed three FAST assessments, each of which probed for the relatedness of stimulus pairs of differing nodal distance. Group- and individual-level analyses broadly demonstrated that relatedness varied as a function of nodal distances in pre-trained stimulus classes. However, results also highlighted some limitations of the FAST at the individual level.

1. Introduction

Since the inception of the Implicit Association Test (IAT; Greenwald et al., 1998), mainstream psychology has seen a wealth of research into the utility of indirect measurement procedures for assessing associations between verbal or other stimuli, usually in socially relevant contexts (e.g. race, gender). Although these procedures originate within the social-cognitive paradigm as measures of associations in memory (Greenwald and Banaji, 1995; Greenwald and Nosek, 2008), implicit measures have also piqued interest within the behavior-analytic field on the basis of their potential utility in indexing verbal learning histories. Two primary behavior-analytic implicit measures have been developed to date: the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2010), and the Function Acquisition Speed Test (FAST; O'Reilly, et al., 2012). While the IRAP has been used almost solely in the context of real-world stimulus relations (with a few exceptions; e.g., Hughes and Barnes-Holmes, 2011; Bortoloti and de Rose, 2011), research using the FAST has been of the "ground-up" variety, with several experimental studies having now been conducted to provide an empirical basis for the interpretation of scores arising from that procedure. That is, the relatedness[1] of stimuli in the FAST scores appears to vary based on the pretraining of stimulus classes via directly trained relations (O'Reilly et al., 2012), derived relations (O'Reilly et al., 2013), across stimulus equivalence relations established using varying amounts of baseline training (Cummins et al., 2018a), as well as experimental manipulations of real world learning histories (Cummins et al., 2018b). These basic studies provide a basis for interpreting the validity of FAST scores when interpreting effects based on real-world stimulus relations (e.g., Cartwright et al., 2016; Roche et al., 2012).

The conceptualization of the FAST is centered on the fact that establishing a functional stimulus class is achieved more readily when the stimuli used already participate in a distinct but compatible functional or equivalence class (Gavin et al., 2008; Roche et al., 2012; Tyndall et al., 2004; see also Urcuioli, 2013). Learning is slower when the stimuli have not been previously related or participate in incompatible functional stimulus or equivalence classes (see O'Reilly et al., 2013, for an empirical example and a brief review of the relevant literature). The FAST utilizes this phenomenon to assess the pre-established learning

---

* Corresponding author.
E-mail address: jamie.cummins@ugent.be (J. Cummins).

[1] Though relatedness may be defined along a number dimensions, research to date has used the term relatedness to refer primarily to the functional equivalence of stimuli. We continue this tradition here.

https://doi.org/10.1016/j.beproc.2020.104179
Received 16 April 2020; Received in revised form 8 June 2020; Accepted 8 June 2020

histories of participants towards stimuli. The FAST consists of two blocks: a ‘consistent’ block, wherein the reinforced response patterns are consistent with the assumed learning history of the participant, and an ‘inconsistent’ block, where the reinforced response patterns are inconsistent with the assumed history of the participant. Participants learn how to respond based on this feedback. Participants’ response fluency increases in each block as they complete trials, but at different rates across the two blocks.

To use a concrete example, the consistent block of a FAST assessing learning histories related to race might involve pressing the ‘z’ key whenever a black face or a negative word is presented and pressing the ‘m’ key whenever a white face or a positive word is presented. The inconsistent block, on the other hand, would require participants to press ‘z’ for black faces and positive words, and press ‘m’ for white faces and negative words. The difference in the fluency of responding on each block provides an index of the direction and magnitude of the learning history of the participant with respect to the specific stimulus relations under analysis. The larger the difference in response fluency across the blocks, the larger the difference in the pre-established fluencies of the two sets of functional or equivalence classes. Where there is no history of contact with the relevant stimuli, there should be no difference in the response fluencies recorded across the blocks (O’Reilly et al., 2012, 2013).

Measuring the variation of stimulus relatedness within different assessment procedures as a function of training has been a stated goal of stimulus equivalence researchers for some time (Bentall et al., 1999; Bortoloti and de Rose, 2009; Moss-Lourenco and Fields, 2011; Sidman et al., 1985). Given its potential for quantifying differences in the verbal learning histories of participants, the FAST may have utility in this regard. In one recent study (Cummins et al., 2018a) scores in the FAST were shown to vary in accordance with the experimenter-manipulated learning histories across six degrees of training. The relatedness of stimuli in the FAST increased as a direct function of increasing amounts of training.

It should be noted that the methodology employed within the Cummins et al. study differed in one salient aspect compared to how learning histories are typically manipulated in stimulus equivalence studies. Specifically, the aforementioned study manipulated learning histories through the use of differing iterations of equivalence training of two different stimulus classes (referred to herein as ‘between-class relatedness’). While this method has been used before in at least one other study (Bortoloti et al., 2013), stimulus relatedness can differ in assessment procedures as a function of the *within* class nodal distance of equivalence class members, referred to herein as ‘within-class relatedness’ (Amd et al., 2018; Bortoloti and de Rose, 2011; Doran and Fields, 2013; Rehfeldt and Dymond, 2005). Nodal distance specifically refers to the number of stimuli intervening between two stimuli in a linearly related set. For example, A-B relations trained directly have a nodal distance of 0, while A-C equivalence relations trained via a B stimulus have a nodal distance of 1.

A variety of other procedures have already been developed within behavior analysis to evaluate the impact of different learning histories on stimulus relatedness. Fields et al. (1995), for example, utilized dual-option function transfer tests. The authors trained participants in 5-member equivalence classes (A-B-C-D-E) using linear MTS training. After this, they trained A to serve as an SD for one response, and E to serve as an SD for another, incompatible response. They observed that the extent to which the response function of the A/E stimulus transferred to the related stimuli differed as a function of nodal distance. This finding has been frequently replicated (Arntzen et al., 2016; Fields, 2015; Fields et al., 2012; Mizael et al., 2016; Moss-Lourenco and Fields, 2011). Bentall et al. (1999) provided another method, observing that response latencies within matching-to-sample testing for equivalence relations varied based on the nodal distance between the sample and comparison being tested. Other methods have also been used, such as post-class formation preference tests (Moss-Lourenco and Fields, 2011), and self-reports such as semantic differentials (Bortoloti and de Rose, 2009). In short, differential nodal distance affects the relatedness of stimuli in a variety of other assessment contexts.

Relative to these measures, the FAST may provide some unique advantages. Firstly, aside from Bentall et al.’s method, none of the currently used measures incorporate response times in defining relatedness within their procedures, in spite of the fact that response times to stimulus relations can provide useful information about the nature of how stimuli are related. By contrast, the FAST amalgamates both response time and accuracy within a single score (the FAST score; see Method for more detail), and so captures both the accuracy and latency dimensions of the effect of relatedness on relational responding. Additionally, because the FAST assesses relatedness in blocks that are both consistent and inconsistent with the learning history of the participant, this offers an opportunity to examine relatedness from two perspectives: *how easy* it is to respond in the learning-consistent block, and *how difficult* it is to respond in the learning-inconsistent block. Although the FAST score amalgamates these two blocks, they can also be examined separately. Additionally, given that different assessment formats can produce widely varying results in assessing stimulus equivalence responding (Doughty and Soydan, 2019), the FAST at the very least represents a novel context which can be used to assess the generalizability of stimulus equivalence or relatedness effects found in other assessment contexts.

For these reasons, the FAST represents a procedure which could make a welcome addition to the toolbox of the equivalence researcher. However, it is not clear whether within-class manipulations will impact the relatedness of stimuli in the FAST in the way that between-class manipulations do (although it is noteworthy that the transfer of stimulus functions within classes varies based on the extent of both training iterations and nodal distance in other assessment contexts; Amd and Roche, 2015; Bortoloti et al., 2013). The suitability of the FAST to produce differential stimulus relatedness effects based on within-class nodal distance manipulations, therefore, requires investigation.

In the current experiment, two 4-member stimulus classes were trained using a linear matching-to-sample (MTS) protocol. The nodal distance between stimulus class members was either 0-nodes (i.e., A-B) 1-node (i.e., A-C) or 2-nodes (i.e., A-D). Following training, all participants completed three different FASTs, with each FAST assessing the relatedness of stimulus pairs of varying nodal distance. It was expected that FAST scores would differ in accordance with the degree of nodal distance between the test stimuli. That is, it was expected that 0-node FASTs would yield the largest effect, followed by the 1-node FAST, followed by the 2-node FAST. Additionally, we expected that relatedness assessed in the consistent FAST block would be inversely related to nodal distance, while relatedness assessed in the inconsistent block would be directly related to nodal distance. We investigated these expectations using both the Frequentist statistical methods commonly used in implicit measures research, as well as the individual-level analyses commonly used in stimulus equivalence research.

2. Method

2.1. Participants

The current experiment was approved by the Maynooth University Research Ethics Committee. The participants within this study ($n = 16$) consisted of Caucasian, Irish undergraduate students attending Maynooth University. Participants were recruited through the use of a University participant volunteer pool and they received no remuneration for their participation. All participants had normal or corrected-to-normal vision and were not afflicted by any condition that may impair performance in tasks requiring sustained attention or learning. All participants completed matching-to-sample (MTS) training, followed by three FASTs in varying order (which individually measured the relatedness of either A-B, A-C, or A-D stimulus pairs), and finally, an MTS testing phase. Nine participants identified as female, while the remaining seven identified as male. The participants had a mean age of 20.94 years (*SD* = 1.73 years).

2.2. Apparatus

The experimental procedure was administered in a small, quiet research/study room (5′ X 5′ approx.) in Maynooth University. All participants engaged in all procedures on a 13″ *Apple MacBook* with a screen resolution of 1024 × 768 pixels. The MTS training and testing procedures were delivered using software created for this research using the experiment generation software PsyScope X (Cohen et al., 1993), while the FAST procedure was delivered via proprietary software produced using *Livecode.* All responses consisted of keyboard button presses or mouse-clicks, and all responses and their timings were recorded by the software programs. All stimuli in the current experiment consisted of three-letter, mono-syllabic nonsense words.

2.3. Procedure

2.3.1. General sequence

At the commencement of the experiment, participants were asked to sign a consent form and were informed that they were free to withdraw from the experiment at any time. All participants completed a matching-to-sample training procedure which trained two 4-member stimulus classes using three-letter nonsense syllables. Following completion of the training, all participants were then required to complete three FASTs. Each FAST used a pair of stimuli from each of the trained classes. The pairs of stimuli differed across the three FASTs in terms of the nodal distance between the stimuli (i.e., in the 0-node FAST, A-B relations were tested; in the 1-node FAST, A-C relations were tested; in the 2-node FAST, A-D relations were tested). For all participants, each FAST took no more than six minutes to complete. Following the completion of the final FAST, participants completed a final MTS testing phase, which tested for responding consistent with B-A, C-A, and D-A stimulus relations.

2.3.2. Stimulus relation training

The current experiment employed a linear matching-to-sample procedure which was designed to establish two 4-member equivalence classes for participants. These classes were first trained by each individual component relation. That is, A1-B1/A2-B2 relations were first trained to a criterion of 15/16 correct responses in a block (16 was chosen as this represented 4 cycles of each of the 2 trial configurations). Then, B1-C1/B2-C2 relations were trained to an identical criterion as the A-B condition. Following this, C1-D1/C2-D2 relations were also trained to the same criterion. When criterion for each of these stimulus relations was met, a training block incorporating trial types from all three prior training blocks was administered. That is, a block presented all of the aforementioned trials in a single block, with each trial cycled three times per block. Participants were required to achieve a criterion of 17/18 correct in this block (18 was chosen as this represented 3 cycles of each of the 6 relevant trial configurations). The predicted emergent classes will be referred to as A1-B1-C1-D1 and A2-B2-C2-D2. The arbitrary stimulus exemplars used in the current experiment consisted of CUG, VEK, JOM, KAH, MAU, ZID, LER, and YOH, with the former 4 constituting the first stimulus class, and the latter constituting the second. The membership of the stimuli as A, B, C, or D stimuli was varied across participants. The following instructions were presented to participants at the beginning of the MTS training phase:

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. More stages will follow after this, with the same task using different words each time. During all stages the computer will provide you with feedback on your performance. Use this feedback to learn how to respond. 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 to begin the Experiment.

On all trials sample stimuli appeared at the center-top of the screen in emboldened size 48 Times font, with two comparison stimuli appearing in the lower left and right corners of the screen one second following the presentation of the sample. There was no time limit on responding. In all blocks of the training procedure, each response was followed by the presentation of feedback in the form of the words ‘correct’ or ‘wrong’ appearing on the computer screen for 1 s. The feedback was accompanied by a brief auditory stimulus; a ‘beep’ sound for correct responses, or a deeper ‘buzz’ sound for incorrect responses. Following this feedback, each trial was proceeded by a 1 s intertrial interval.

As mentioned, the MTS training consisted of two main stages: the sequential phase, and the simultaneous phase. Two stimulus relations were trained during the first block of the sequential phase (A1-B1, A2-B2). In this block, the sample stimulus was always an A stimulus, and the comparison stimuli were always the B1 and B2 stimuli. Following the achievement of criterion, participants were provided with a short break (up to one minute), and upon their prompting (via a mouse click), moved on to the second block of training. This second block involved the training of two further stimulus relations (B1-C1, B2-C2). In this block, the sample stimulus was always a B stimulus, and the comparison stimuli were always the C1 and C2 stimuli. When the participant reached criterion here, they moved on to the third block, which involved identical configurations to the previous two blocks, but trained C1-D1 and C2-D2 stimulus relations. When criterion was reached after this third block, the participant moved on to the simultaneous phase of the training. This involved the further training of all six of the previous stimulus relations (i.e., A1-B1, A2-B2, B1-C1, B2-C2, C1-D1, and C2-D2) in simultaneity. The trial configurations were the same as those in the sequential phase in terms of sample and comparison presentation.

In the sequential phase, each trial type was presented eight times in total per training block, leading to a total of 16 trials per block of training. Participants were required to reach a criterion minimum of fifteen out of sixteen correct responses in a block in order to progress to the next relevant block. If participants failed to reach criterion on eight consecutive training blocks, then the participant was simply moved on to the next block of training; this was in order to prevent participants from experiencing boredom due to the repetitiveness of a specific block, and also due to the fact that participants would have another opportunity to demonstrate responding to sufficient criterion in the simultaneous phase.

Once participants completed each of the three sequential blocks of training, they immediately moved on to the simultaneous training phase. In this phase, participants were tested on each of the six stimulus relations presented in the previous blocks. That is, six trial types in this phase were present: A1-B1 (B2), A2-B2 (B1), B1-C1 (C2), B2-C2 (C1), C1-D1 (D2), and C2-D2 (D1), where the stimuli in parentheses represent incorrect choices. Each of the six trial types was presented in a quasi-random order, with each trial presented once in a cycle of six trials. Again, no trial type was presented more than twice in succession. Each trial type was presented three times in each iteration of this phase, leading to each simultaneous block consisting of 18 trials. The criterion for passing this phase was seventeen out of eighteen responses correct on a single block. If the criterion was not achieved, the block recycled. Like the sequential phase, the simultaneous block was allowed to cycle eight times. If eight cycles passed and the participant had not passed the block, the procedure terminated. A participant was **not** considered to have failed the training if they failed to reach criterion in the sequential stage but achieved criterion in the simultaneous stage. However, failure to achieve criterion in the simultaneous stage was considered to be a failure of training. If this occurred, participants would be thanked for their time, debriefed, and their data omitted from the study. However, all participants passed the simultaneous phase on training within the eight cycles. Note that no testing for emergent symmetrical/equivalence relations was done in this phase.

2.3.3. Function acquisition speed test

Immediately following the training phase, participants were required to complete three Function Acquisition Speed Tests (FASTs) in a counter-balanced order. In each FAST there were four experimental stimuli used. In the 0-node FAST, the A1-B1/A2-B2 stimuli from the MTS training were used; in the 1-node FAST, the A1-C1/A2-C2 stimuli were used; and in the 2-node FAST, the A1-D1/A2-D2 stimuli were used. Each FAST consisted of two blocks (‘consistent’ and ‘inconsistent’), both of which contained 50 trials. The order of presentation of the two blocks was randomized in all administrations. In all blocks, each trial involved the presentation of a single stimulus in 32-point font in the center of the computer screen. Upon seeing the stimulus, participants were required to respond using either the ‘z’ or ‘m’ keys on the computer keyboard. In each block, two stimuli served as S_Ds for the ‘z’ response, while two other stimuli served as S_Ds for the ‘m’ response.

Importantly, in the ‘consistent’ blocks, the S_Ds for the ‘z’ response, and the S_Ds for the ‘m’ response, were consistent with the trained stimulus classes. For example, in the consistent block of the 0-node FAST, A1 and B1 were discriminative of a ‘z’ response, while A2 and B2 were discriminative of an ‘m’ response. In this sense, participants were required to learn that A1 and B1 participated in a shared functional response class, and that A2 and B2 participated in another shared functional response class. In the ‘inconsistent’ block, a converse configuration was arranged; that is, A1 and B2 were discriminative for the ‘z’ response, and A2 and B1 were discriminative for the ‘m’ response. The key manipulation here was that in the inconsistent block, participants were required to respond in a way that was inconsistent with the experimentally-produced learning history, while in the consistent block participants were required to respond in a way that was consistent with this history. It was expected that learning to respond in a way that was consistent with the prior history should be easier (i.e., learning should be more fluent) than learning to respond in a way that was inconsistent with the prior learning history.

A key press by the participant initiated the FAST procedure, after the following minimal instructions were presented and read on-screen:

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

This was followed by an intertrial interval (ITI) of 500 ms, after which the first trial began. There was a 3000 ms limited hold on each trial. If a response was emitted within 3000 ms then the stimulus was removed from the screen, and feedback was presented as appropriate. If the participant did not respond within the limited hold, then the stimulus was removed from the screen and corrective feedback for an erroneous response (i.e., ‘wrong’) was presented and a response time of 3000 ms was recorded. The stimuli were presented in a quasi-random order, with each stimulus presented once in a given cycle of four trials. Each block of the FAST consisted of 12.5 cycles, with two stimuli presented an extra time to complete the block of 50 trials.

The experimental stimuli used in each of the three FASTs varied in terms of the nodal distance between the relevant procedural stimuli. In the 0-node FAST, the stimulus pairs A1-B1 and A2-B2 consisted of the consistent block, while the alternative configuration (i.e., A1-B2, A2-B1) consisted of the inconsistent block. In the 1-node FAST, the relevant consistent stimulus pairs were A1-C1 and A2-C2, while in the 2-node FAST, the consistent stimulus pairs were A1-D1 and A2-D2. The FASTs were presented in a counter-balanced order across participants in terms of the six possible presentation orders.

Scores in the FAST were calculated by first fitting a regression line to the cumulative learning curve of each FAST block (with cumulative response time on the X axis, and accuracy on the Y axis). The slope of this regression line is referred to as the block-slope score. A larger block-slope score indicates a steeper regression line, which in turn indicates greater fluency of responding in the given block. In line with previous studies, the block-slope score of the inconsistent block was subtracted from the block-slope of the consistent block to produce the FAST effect size, which represents the difference in fluency of responding between the two FAST blocks. A positive FAST score indicates more fluent responding on the consistent block compared to the inconsistent block. Conversely, a negative FAST score indicates that responding was more fluent on the inconsistent compared to the consistent block.

2.3.4. Stimulus relation testing

Upon completion of the three FAST procedures, participants were required immediately after to complete a partial test for stimulus equivalence. This testing phase consisted of the same MTS format as in the training phase, but with three salient differences. Firstly, no feedback was presented to participants following their responses on each trial. Secondly, this testing phase was of a fixed trial length: sixty trials for all participants, with all relevant stimulus relations tested simultaneously. Finally, the samples and comparisons in the MTS testing phase differed from the training phase. The trial configurations consisted of A1-B1 (B2), A2-B2 (B1), B1-A1 (A2), B2-A2 (A1), A1-C1 (C2), A2-C2 (C1), C1-A1 (A2), A2-C2 (C1), A1-D1 (D2), A2-D2 (D1), D1-A1 (A2), and D2-A2 (A1), with comparison stimuli to the right of the dash, and the erroneous response option parenthesized. These twelve trial configurations were presented five times each in quasi-random order, with each trial presented once per twelve trials, and no trial presented twice in a row at any stage. Thus, ten of the twenty trials for A-B stimulus relations involved using the A stimulus as sample, while ten involved using the B stimulus as sample (this was also true of A-C and A-D relations). Upon completion of the sixty trials, the procedure terminated, and the participant was informed that the experiment was complete.

2.3.5. Measurement

Data for the MTS training procedure were recorded for the number of trials of training required to reach criterion on each of the sequential blocks, as well as on the simultaneous block. The mean response times of each block were also recorded. For the FAST, the primary datum of interest was that of FAST score (response fluency differential across blocks). Data from the MTS testing phase was taken in terms of the number of correct responses out of twenty that participants achieved across the probes for each of the three stimulus relations (i.e., A-B, A-C, and A-D).

3. Results

3.1. Group level

3.1.1. Descriptive statistics

3.1.1.1. Stimulus relation training. Initial descriptive analyses discerned no differences in the number of training cycles required, nor the mean response time required, to pass sequential MTS training across each of the three sets of stimulus relations. All participants passed the simultaneous phase of training within the eight-cycle limit imposed. One participant, Participant 10, failed to achieve criterion in the A-B sequential training block, but achieved criterion in the B-C and C-D sequential training, as well as the pooled simultaneous training (see Table 1).

3.1.1.2. Function acquisition speed tests. Descriptive analyses showed that, as expected, FAST scores generally decreased as the nodal distance

Table 1
The number of block cycles needed by, and mean response times of, each participant in the sequential and simultaneous MTS training stages.

	MTS Training Phase							
	A-B		B-C		C–D		Simultaneous	
Participant	Cycles	RT	Cycles	RT	Cycles	RT	Cycles	RT
1	1	1167.25	1	1087.94	1	1266.38	3	1628.78
2	2	2337.22	2	1655.34	1	1623.75	7	1935.97
3	1	1286.38	2	1024.00	1	1161.31	1	1358.22
4	1	1529.50	1	1608.81	1	1554.75	1	2078.72
5	1	1493.31	1	1856.94	1	2027.69	1	2905.28
6	1	1564.44	1	1526.25	1	1400.56	2	1672.69
7	1	1976.75	1	2669.00	1	2205.81	1	2049.22
8	2	2985.88	1	1316.44	2	1479.81	1	1510.06
9	2	2451.63	2	1895.09	1	1678.56	3	3360.11
10	8	774.90	2	1227.97	2	1387.33	2	1335.11
11	3	1632.26	1	2727.19	2	3465.38	2	3678.56
12	1	1683.38	1	1310.19	1	1834.56	2	2629.36
13	1	2347.56	1	1786.13	2	1741.84	1	2652.83
14	2	1610.84	1	1473.75	1	1552.63	4	2115.24
15	4*	1228.79	4*	1337.41	4*	1257.44	5	1570.08
16	1	2276.81	3	1622.38	1	1191.94	4	2470.79

* Due to a technical fault, Participant 15 continued sequential training after criterion was met. Visual analysis of the data indicated that the participant had achieved criterion after two cycles of A-B and C-D, and after one cycle of B-C.

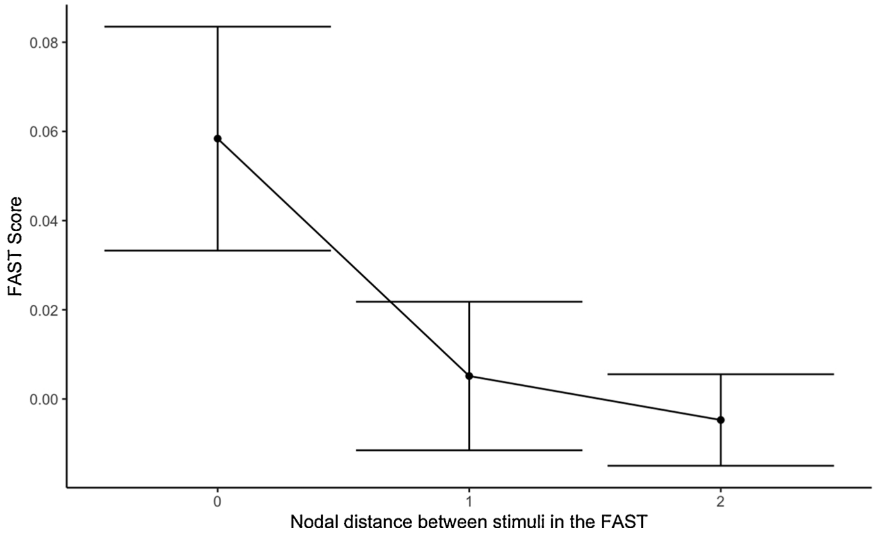


Fig. 1. Mean FAST score at each level of nodal distance. Scores below zero represent unexpected (untrained) response patterns. Error bars refer to +/- one standard error of the mean.

of the FAST stimuli increased. Fig. 1 illustrates the aggregated FAST scores for each of the three FASTs. Fig. 2 shows the changes in FAST scores on each of the FASTs at the individual participant level. Table 2 displays the means for FAST scores for each participant's three FASTs.

3.1.1.3. Stimulus relation testing. As per typical criteria in stimulus equivalence studies (e.g., Fields and Moss, 2007), participants were considered to have passed the test for the relevant stimulus relation if they achieved eighteen out of twenty correct responses for each of the three relevant stimulus relations (i.e., 90 % correct). According to these criteria, 10 out of 16 participants passed the test for A-B stimulus relations, 6 out of 16 participants passed the test for A-C stimulus relations, and 8 out of 16 passed for A-D relations.

We subsequently examined the extent to which passing/failing the stimulus relation testing corresponded with the presence of a positive/negative FAST score. As is apparent in in Table 3, the strength of A-B and A-C relations, as measured by the FAST, varied more than they did according to the MTS testing. However, A-D relations varied more in their degess of relatedness during the MTS testing than during the FAST.

3.1.1.4. Inferential statistics. Inferential analyses were run in order to determine whether there was any significant effect of nodal distance on FAST scores in each of the three FASTs. We used a one-way repeated-measures ANOVA to examine if such a change was present. The assumption of normality was maintained for this test, D(45) = 0.14, *p* = 0.25. As expected, a significant change was noted for FAST scores across the 0-node (*M* = .06, SD = .10), 1-node, (*M* = 0.01, SD = .07), and 2-node (*M* = -0.01, SD = .04) FASTs, (*F* (2, 14) = 5.21, *p* = .011; η_G^2 = .132). We next conducted post-hoc t-tests to unpack this interaction. Kolmogorov-Smirnov tests for each of the pairs (i.e., 0-node and 1-node, 0-node and 2-node, and 1-node and 2-node) indicated that the distribution of FAST scores differed significantly from the normal distribution for each pair (all *p*s < .001). We therefore conducted a series of Wilcoxon Signed-Rank tests. Additionally, we conducted Holm-Bonferroni correction on the p-values for these tests to correct for multiple comparisons (Aickin and Gensler, 1996). There was a significant difference between both the 0-node and 1-node conditions

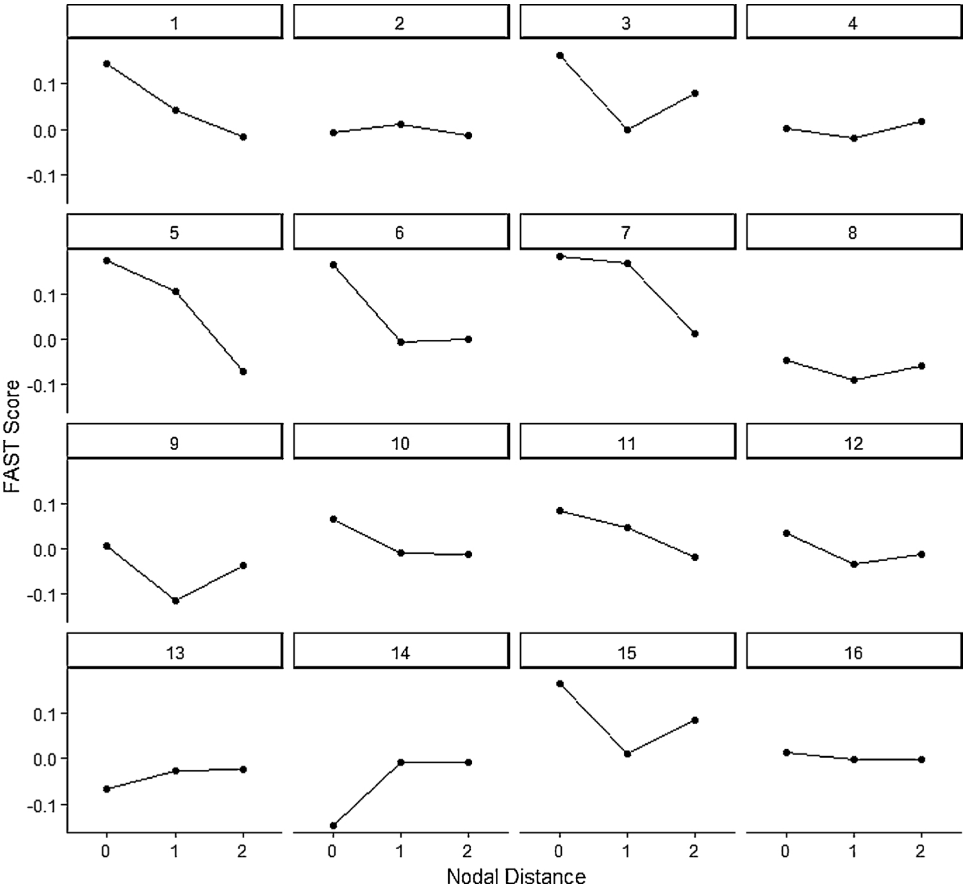


Fig. 2. FAST scores for each participant at each level of nodal distance. Scores below zero represent unexpected (untrained) response patterns.

Table 2
FAST scores for each of the three FASTs. A positive score represents a score in the expected direction.

	FAST Score		
Participant	0-Node FAST	1-Node FAST	2-Node FAST
1	.14	.04	−.02
2	−.01	.01	−.01
3	.16	.00	.08
4	.00	−.02	.02
5	.17	.11	−.07
6	.16	−.01	.00
7	.18	.17	.01
8	−.05	−.09	−.06
9	.01	−.11	−.04
10	.07	−.01	−.01
11	.08	.05	−.02
12	.04	−.03	−.01
13	−.07	−.03	−.02
14	−.15	−.01	−.01
15	.16	.01	.08
16	.01	−.00	−.00

Table 3
Relative pass/failure of each participant on the FAST and MTS for each of the 3 conditions. Passing in the FAST refers to a training-consistent FAST score (i.e., > 0), while failing refers to a training-inconsistent FAST score (i.e., < 0). Passing in the MTS refers to achieving at least 18 out 20 correct responses. Scores marked with an asterisk denote instances in which there was agreement across the two measures.

	Stimulus relation					
	A-B		A-C		A-D	
Participant	FAST	MTS	FAST	MTS	FAST	MTS
1	PASS	FAIL	PASS	FAIL	FAIL	PASS
2	FAIL	PASS	PASS	FAIL	FAIL	PASS
3	PASS*	PASS*	PASS*	PASS*	PASS*	PASS*
4	PASS*	PASS*	FAIL	PASS	PASS*	PASS*
5	PASS*	PASS*	PASS*	PASS*	FAIL	PASS
6	PASS*	PASS*	FAIL*	FAIL*	PASS	FAIL
7	PASS*	PASS*	PASS*	PASS*	PASS*	PASS*
8	FAIL*	FAIL*	FAIL*	FAIL*	FAIL*	FAIL*
9	PASS*	PASS*	FAIL*	FAIL*	FAIL*	FAIL*
10	PASS	FAIL	FAIL*	FAIL*	FAIL*	FAIL*
11	PASS*	PASS*	PASS*	PASS*	FAIL	PASS
12	PASS	FAIL	FAIL*	FAIL*	FAIL*	FAIL*
13	FAIL	PASS	FAIL	PASS	FAIL	PASS
14	FAIL*	FAIL*	FAIL*	FAIL*	FAIL*	FAIL*
15	PASS	FAIL	PASS	FAIL	PASS	FAIL
16	PASS*	PASS*	FAIL*	FAIL*	FAIL*	FAIL*

(Z = 2.42, corrected p = .033) and the 0-node and 2-node conditions (Z = 2.54, corrected p = .033). However, we did not find a significant difference between the 1-node and 2-node conditions (Z = 0.13, corrected p = .90).

A further analysis was run in order to investigate whether there was any interaction between the FAST blocks and nodal distance in terms of scores for each FAST. A 2 (FAST block: consistent or inconsistent) x 3 (nodal distance: 0-node, 1-node, or 2-node) within-participants ANOVA was run in order to investigate this. The assumption of normality was maintained, D(90) = 0.10, $p = 0.207$. A significant interaction effect was found between FAST block and nodal distance (F (2, 30) = 5.21, $p = .0115$; $\eta^2_G = 0.058$; see Figs. 3 and 4 for illustration of this trend), indicating that response fluency was differentially affected by nodal distance across the two FAST blocks. In order to further investigate this

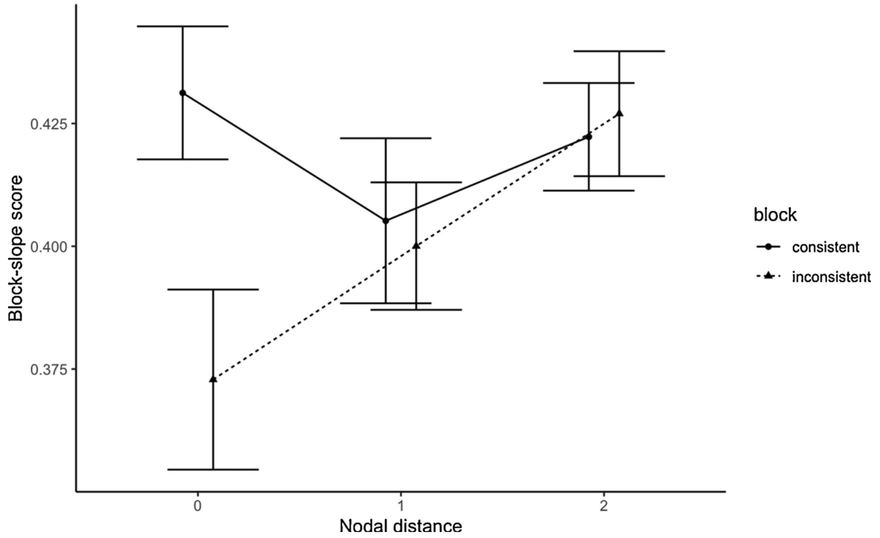


Fig. 3. Differences in the mean block-slope scores at differing levels of nodal distance. Error bars refer to +/- one standard error of the mean.

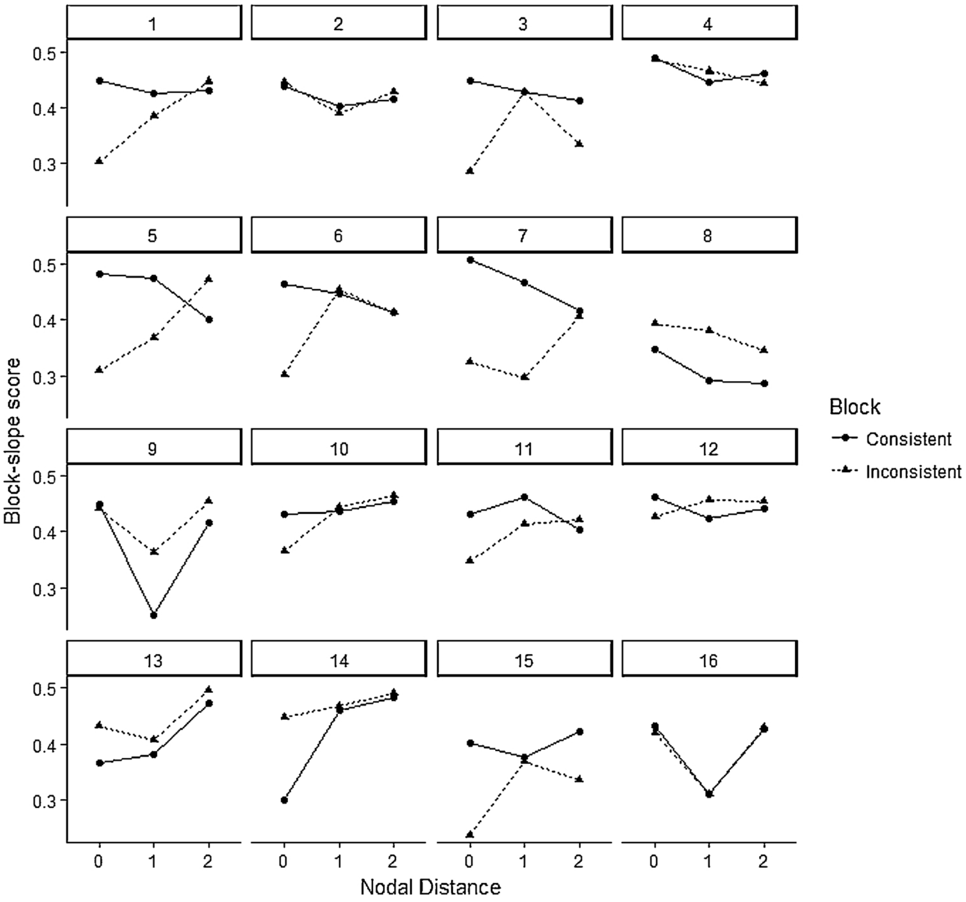


Fig. 4. Changes in the block-slope scores of each individual participant at differing levels of nodal distance.

interaction in terms of whether one of the two FAST blocks better conformed to the expected linear trends, we ran two additional one-way ANOVAs: one for block-slope scores in the consistent block only, and one for block-slope scores in the inconsistent block only. For the inconsistent block the assumption of normality was maintained (D (45) = 0.09, $p = 0.774$), and block-slope scores increased as a function of increasing nodal distance as expected, $F(2, 30) = 4.95$, $p = .0139$; $\eta^2_G = 0.128$. The assumption of normality was also maintained for the consistent block (D(45) = 0.19, $p = .057$), but block-slope scores did not vary with differential nodal distances, $F(2, 30) = 0.88$, $p = 0.323$; $\eta^2_G = .038$.

A final analysis was run in order to determine whether any order effects were present for block presentation in each of the three FASTs for FAST score. As such, three independent sample t-tests were run. No significant effect of FAST block order was found on the 0-node FAST for FAST score ($t(14) = 0.99$, $p = 0.337$, Cohen’s d = .50, 95 % CI [-0.51, 1.50]). For the 1-node FAST, there was also no significant impact of block order on FAST score ($t(14) = 0.01$, $p = 0.992$, Cohen’s d = 0.01, 95 % CI [-1.01, 1.02]). This was also the case for the 2-node FAST (t (14) = 0.05, $p = 0.961$, Cohen's d = 0.03, 95 % CI [-0.99, 1.04]).

3.1.1.5. Individual level. We expected that individual participants would show decreasing FAST scores as a function of increasing nodal distance. Overall, 10 participants showed trends in responding which broadly aligned with our expectations. The exact expected trend was seen for 3 participants: 1, 5, and 11. Additionally, 7 participants (3, 6, 7, 9, 10, 12, and 15) showed a broadly consistent trend, with slight variations: 3 participants (6, 10, and 12) differed in responding in the 0-node FAST compared to the other two conditions, though these two conditions did not differ, and 4 participants (3, 7, 9, and 15) showed the expected decline in FAST scores from the 0-node condition to the 2-node condition, but with irregular 1-node condition scores (Participants 3, 9, and 15 showed 1-node scores lower than the 2-node scores, while Participant 7 showed a 1-node score which was similar to the 0-node score). Six participants showed trends in responding which deviated from our expectations. Specifically, five participants (2, 4, 8, 13, and 16) showed approximately identical FAST scores across all 3 nodal distances, while the remaining participant (Participant 14) showed identical FAST scores for the 1- and 2-node conditions, but showed a much lower FAST score in the 0-node condition.

We also expected that consistent block-slope scores would decrease, and inconsistent block-slope scores would increase, as a function of increasing nodal distance. The expected consistent block trend was seen for 10 participants. The trends of 3 participants (3, 6, and 7) exactly matched our expectations. In addition, 7 participants (2, 4, 5, 8, 9, 11, and 12) showed a broadly similar pattern, with some deviation: these participants showed the expected decline from 0-node to 2-node, but with variation on the position of 1-node scores (lower than both 0-node and 2-node scores for participants 4 and 9; identical to the 0-node score for Participant 5; identical to the 2-node score for Participants 2, 8, and 12; and higher than both the 0- and 2-node scores for Participant 11). The remaining 6 participants (1, 10, 13, 14, 15, and 16) did not respond as expected. Participant 1 showed identical slope scores on the consistent block for all 3 nodal distances, while Participants 10, 13, and 14 showed opposite trends in the block than what was expected. In addition, Participants 15 and 16 both showed depressed block-slopes in the 1-node condition relative to the other conditions, with Participant 15 also showing an unexpected larger block-slope score for the 2-node condition compared to the 0-node condition.

For the inconsistent block, we expected that block-slope scores would increase as a function of nodal distance: this trend was demonstrated by 12 participants. Five participants (1, 5, 10, 11, and 14) behaved wholly in line with our expectations. An additional 6 participants (3, 6, 7, 9, 12, 13, and 15) showed the expected trend from the 0-node to the 2-node conditions, but with variable performance on the 1-node conditions: for Participants 3, 6, and 15, 1-node scores were higher than in the 2-node condition; for Participants 7, 13, and 9, 1-node scores were lower than in the 0-node condition; and for Participant 12, the block-slope score for the 1-node condition was approximately equal to that of the 2-node condition. The remaining 4 participants (2, 4, 8, and 16) did not show the expected block-slope scores in the inconsistent block. For Participants 2 and 16, the 0-node and 2-node conditions showed equal block-slope scores, with lower 1-node scores. Additionally, Participants 4 and 8 showed exactly opposite patterns to what was expected: block-slope scores *decreased* as a function of nodal distance.

3.1.1.6. Summary. At the group level, a decrease in FAST scores (and mean response times) was seen as the nodal distance between FAST stimuli increased. However, upon examining this main effect further, significant differences across conditions were observed across the 0-node and 1-node conditions, and the 0-node and 2-node conditions, but not the 1-node and 2-node conditions. There was an interaction effect between FAST block and nodal distance for block-slope scores, with consistent block-slope scores being unaffected by nodal distance, and inconsistent block-slope scores increasing as a function of nodal distance. As well as this, no confounding effect of FAST block order was found. At the individual level, 10 out of 16 participants showed overall FAST score trends which were broadly consistent with our expectations. This was the case also for 10 participants in terms of consistent block-slope scores, and 12 participants for inconsistent block-slope scores.

4. Discussion

In this experiment we sought to investigate the effect of differential nodal distance on the relatedness of stimuli in the Function Acquisition Speed Test. Each participant was trained in the formation of two 4-member stimulus equivalence classes. The participants then completed three FASTs in a counter-balanced order. Each FAST consisted of the formerly trained stimuli but differed in terms of their nodal distance (i.e., A-B, A-C, or A-D relations). It was expected that FAST scores would decrease across the FASTs as a function of increasing nodal distance between the experimental stimuli.

4.1. Effects at the group level

The finding of a significant effect of nodal distance on FAST scores in the ANOVA was in line with our expectations. However, in post-hoc analyses, a significant difference was seen only between the 0-node and 2-node FASTs and the 0-node and 1-node FASTs, but not between the 1-node and 2-node FASTs. Even at the descriptive level, the difference between the 1-node and 2-node conditions was negligible, indicating that stimulus relatedness in the FAST was not affected by differential nodal distances in these derived stimulus relations.

The finding of a significant interaction effect between FAST block and nodal distance allows for further dissection of the above finding. Slope scores in the inconsistent block of the FAST varied in accordance with what was expected: inconsistent block slope scores decreased as a function of decreasing nodal distance indicating, in simple terms, that participants found it more difficult to respond in opposition to the established relations when those relations were more strongly established. For the consistent block however, such an effect was not seen: slope scores in this block did not vary as a function of nodal distance. Although unexpected, previous research can contextualize this finding insofar as changes in stimulus relatedness as measured by the FAST have previously been observed to involve changes in the slope of the inconsistent block, and not changes in the consistent block (Cummins et al., 2018a).

In the same sense that more difficult math questions could better discriminate between the performance of a mathematics professor and a high school student, increasing the "difficulty" of the consistent block of the FAST could increase the likelihood that the relatedness of stimuli will also differ in this block. One method of doing this may be to constrain the length of the response window within trials. Such a constraint would demand generally shorter response times, which would serve to increase the rate of errors, thereby reducing fluency. In doing this, stimulus relations of less nodal distance should still of course facilitate the formation of classes more than stimuli of greater nodal distance, but this might then be reflected more evenly across both FAST blocks, rather than in inconsistent block slopes alone. Notably, the length of the response window should not lead to a fluency reduction to the point of eliminating differences across blocks (i.e., through the general truncation of variances; see Cummins et al., 2018a). Determining the optimal length for response windows requires a careful empirical analysis. Nevertheless, at this stage it presents itself as one obvious means by which we might improve the likelihood that stimulus relatedness in the FAST might differ as a function of different learning histories.

4.2. Effects at the individual level

Results of analyses conducted at the level of individual participants were somewhat equivocal. On the one hand, 10 participants showed changes in FAST scores which were in line with what was expected. However, effects for the other 6 participants were more mixed, and one participant in particular (Participant 14) showed effects which were essentially opposite to what was expected. This less-than-perfect individual-level performance may be somewhat expected, given that implicit measures in general are developed for, and employed at the group level (Gawronski and De Houwer, 2014). However, we should also bear in mind that only one of the 16 participants showed a response pattern that was directly in opposition to what we expected. Nevertheless, consistent individual-level performance is ultimately required if the FAST is to be useful for within stimulus equivalence research. In this sense, the current findings do not provide as strong support for the FAST as was expected, but they do corroborate the idea that stimulus relatedness varies in the FAST as a function of nodal distance for most participants. This finding represents a very important step in the development of a functional account of the behavioral processes at work in implicit measures. Such an account has been sorely lacking to date, despite a literature top-heavy with applications of implicit measures within the field (though see Finn et al., 2018 for a recent exception).

One potential explanation for our mixed findings at the individual level could be due to the lack of remuneration participants received, which would reduce the likelihood of some participants providing meaningful data. However, this is unlikely; previous research has shown that participants without remuneration produce meaningful learning effects in the FAST based on the training of equivalence relations (Cummins et al., 2018a; O'Reilly et al., 2013). Additionally, several commonly used procedures for measuring stimulus relatedness are also susceptible to similar individual-level variability. For instance, in Bentall et al. (1999), relatedness differences were best observed at the group level, but varied widely at the individual level. By contrast, measures which utilize discrete (rather than continuous) dependent variables (i.e., post-class formation preference tests, dual-option function transfer tests, and semantic differentials) tend to demonstrate less variability (although this is less true for semantic differentials at the individual-level; see Bortoloti and de Rose, 2009). Post-class formation preference tests and dual-option function transfer tests in particular tend to be highly uniform in their effects, *but only once participants have passed tests of the relevant relations using another procedure.*

The foregoing point is highly salient. Specifically, in our analyses we included all participants, including those who *failed* MTS testing for the trained stimulus relations. Additionally, unlike other studies, we did not recycle participants back into training if they failed MTS testing (Moss-Lourenco and Fields, 2011; Bortoloti et al., 2019). As such, it might be that further training to criterion on MTS testing would have produced greater differential relatedness of stimuli across the different FASTs. When comparing individual performances on the FASTs and MTS testing, we observe some interesting trends. In general, the behavior of an individual in the FAST and in the MTS testing was consistent (i.e., both showed training-consistent or training-inconsistent responding) for 10 out of 16 participants for A-B stimulus relations, 11 out of 16 participants for A-C relations, and 9 out of 16 participants for AD relations (see Table 4). Furthermore, for both the 0-node and 1-node conditions, the relatedness of stimuli in the FAST varied to an even greater extent than MTS testing. For A-B stimulus relations, 12 out of 16 participants demonstrated a positive FAST score, while only 10 participants reached criterion in the MTS. Likewise, for the A-C stimulus relations, 10 out 16 participants demonstrated a positive FAST score, while only 6 showed reached criterion in MTS testing. In contrast, relatedness of stimuli in MTS testing varied more as a function of training than in the FAST for the 2-node relations, with 8 out of 16 reaching criterion in MTS testing, but only 6 out of 16 showing a positive FAST effect. Overall, the measures corroborated in approximately 70 % of cases, and relatedness in the FAST was more sensitive to the trained stimulus relations for both the 0-node and 1-node relations.

The above point should be carefully caveated by the fact that we did not include a control group to compare base rates of MTS testing performances. That is, all participants completed the measures in an identical order: MTS training, followed by the three FASTs, and then MTS testing. As such, participants' performances on the MTS testing may well have been influenced by their previous completion of the FASTs. Given that the FASTs required response patterns which were both consistent *and* inconsistent with learning, it is difficult to know what effect this might have had (e.g., they could "cancel each other out" and have no effect, or influence learning histories in highly unpredictable ways and have a substantial effect). As such, our comparisons between FAST and MTS testing performances are limited by this potential confound. Although participants' mixed performance on the MTS testing can alternatively be argued as being typical for the first round of equivalence testing (i.e., often participants must cycle back through training multiple times before they can correctly emit all equivalence responses), the uncertainty around this represents a weakness of our study which should be addressed in future work. By measuring learning histories with the FAST, we may well perturb those same histories.

4.3. Implications for the (behavior-analytic) use of implicit measures

This work represents an important, yet perhaps not immediately obvious, contribution to the general use of implicit measures within behavioral and experimental psychology. Specifically, the behavior-analytic tradition of individual-level analysis has not generally been utilized in research on implicit measures. While early research using the IRAP presented individual mean latencies for participants, and research using the FAST has formerly presented similar information (Power et al., 2009; O'Reilly et al., 2012), even these studies only provided this information summarily, with little further analysis. In present-day research, reporting on individual-level statistics is essentially non-existent (though see Finn, 2020 for a recent exception). This is somewhat surprising, given that the use of these measures by behavior analysts is increasing at a strikingly high rate (Dymond and May, 2018). In spite of this, the analysis of these measures is done exclusively at the group-level, following on from the social-cognitive tradition of the use of these measures.

Viewing under the hood of group-level statistics is essential for behavior analysts, or indeed any experimental psychologists, who use these implicit measures. In particular, the IRAP has been, and is currently being used in a wide variety of contexts to distinguish groups of the basis of verbal histories. These contexts are as varied as (for example) predicting spider fear (Nicholson and Barnes-Holmes, 2012), sex offender status (Dawson et al., 2009), and anorexia status (Parling et al., 2012). However, these group-level findings do not provide any insight into the specificity of the measure in quantifying verbal histories for any given individual. Indeed, as the present research demonstrates, by presenting both group- and individual-level analyses, group-level statistics can serve to obfuscate individual-level variability and can give the impression that the measure is more effective than it may be in reality (or hide its utility at the individual-level through poor group-level estimation with small sample sizes). This is problematic generally, but particularly for studies which seek to use the relatedness of stimuli in some measure to predict an external outcome (e.g., success of treatment for cocaine dependency; Carpenter et al., 2012).

By ignoring individual variability through the exclusive use of group-level statistics, this can also impede measures in the long-run, even at the group-level. For a measure to be effective in quantifying verbal learning histories, it is important for responding to be solely under the control of the stimulus relations being assessed. However, group-level statistics can also mask the influence of unwanted sources of stimulus control over the behavior of individuals within the measure (for example block order effects, which are known to have an impact on performance in these measures, but the functional nature of this impact is poorly understood). This oversight can reduce the statistical power of the measure to detect effects at the group level, by injecting variability into scores in the measure which are not due to the relations being assessed. In statistical terms, this results in poorer estimations of groups' mean levels of responding. As such, for implicit measures researchers, it is imperative to consider such an individual-level treatment of data. This study represents one of the first efforts to outline this relationship between group-level and individual-level performances in the context of these measures, and as such offers a baseline against which tests can be improved in an effort to bring individual level performances under greater experimental control.

Author's note

Production of this manuscript was part-funded a John and Pat Hume scholarship awarded to JC under the supervision of BR.

CRediT authorship contribution statement

Jamie Cummins: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing - original draft. **Bryan Roche:** Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Software, Supervision, Validation, Writing - review & editing.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.beproc.2020.104179.

References

Aickin, M., Gensler, H., 1996. Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods. Am. J. Public Health 86 (5), 726–728.

Amd, M., Roche, B., 2015. A derived transformation of valence functions across two 8-member comparative relational networks. Psychol. Rec. 65 (3), 523–540. https://doi.org/10.1007/s40732-015-0128-1.

Amd, M., de Oliveira, M., Passarelli, D., Balog, L., de Rose, J., 2018. Effects of orientation and differential reinforcement II: transitivity and transfer across five-member sets. Behav. Processes 150, 8–16. https://doi.org/10.1016/j.beproc.2018.02.012.

Arntzen, E., Granmo, S., Fields, L., 2016. The relation between sorting tests and matching-to-Sample tests in the formation of equivalence classes. Psychol. Rec. 67 (1), 81–96. https://doi.org/10.1007/s40732-016-0209-9.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., Boles, S., 2010. A sketch of the implicit relational assessment procedure (IRAP) and the relational elaboration and coherence (REC) model. Psychol. Rec. 60 (3), 527–542. https://doi.org/10.1007/bf03395726.

Bentall, R., Jones, R., Dickins, D., 1999. Errors and response latencies as a function of nodal distance in 5-member equivalence classes. Psychol. Rec. 49 (1), 93–115. https://doi.org/10.1007/bf03395309.

Bortoloti, R., de Rose, J., 2009. Assessment of the relatedness of equivalent stimuli through a semantic differential. Psychol. Rec. 59 (4), 563–590. https://doi.org/10.1007/bf03395682.

Bortoloti, R., de Rose, J., 2011. An "Orwellian" account of stimulus equivalence. Are some stimuli "more equivalent" than others? Eur. J. Behav. Anal. 12 (1), 121–134. https://doi.org/10.1080/15021149.2011.11434359.

Bortoloti, R., Rodrigues, N., Cortez, M., Pimentel, N., de Rose, J., 2013. Overtraining increases the strength of equivalence relations. Psychol. Neurosci. 6 (3), 357–364. https://doi.org/10.3922/j.psns.2013.3.13.

Bortoloti, R., de Almeida, R.V., de Almeida, J.H., de Rose, J.C., 2019. Emotional faces in symbolic relations: a happiness superiority effect involving the equivalence paradigm. Front. Psychol. 10. https://doi.org/10.3389/fpsyg.2019.00954.

Carpenter, K., Martinez, D., Vadhan, N., Barnes-Holmes, D., Nunes, E., 2012. Measures of attentional Bias and relational responding are associated with behavioral treatment outcome for cocaine dependence. Am. J. Drug Alcohol Abuse 38 (2), 146–154. https://doi.org/10.3109/00952990.2011.643986.

Cartwright, A., Roche, B., Gogarty, M., O'Reilly, A., Stewart, I., 2016. Using a modified Function Acquisition Speed Test (FAST) for assessing implicit gender stereotypes. Psychol. Rec. 66 (2), 223–233. https://doi.org/10.1007/s40732-016-0164-5.

Cohen, J., MacWhinney, B., Flatt, M., Provost, J., 1993. PsyScope: an interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. Behav. Res. Methods Instrum. Comput. 25 (2), 257–271. https://doi.org/10.3758/bf03204507.

Cummins, J., Roche, B., Tyndall, I., Cartwright, A., 2018a. The relationship between differential stimulus relatedness and implicit measure effect sizes. J. Exp. Anal. Behav. 110 (1), 24–38. https://doi.org/10.1002/jeab.437.

Cummins, J., Tyndall, I., Curtis, A., Roche, B., 2018b. The function acquisition speed test (FAST) as a measure of verbal stimulus relations in the context of condom use. Psychol. Rec. 69 (1), 107–115. https://doi.org/10.1007/s40732-018-0321-0.

Dawson, D., Barnes-Holmes, D., Gresswell, D., Hart, A., Gore, N., 2009. Assessing the implicit beliefs of sexual offenders using the implicit relational assessment procedure. Sex. Abus. A J. Res. Treat. 21 (1), 57–75. https://doi.org/10.1177/1079063208326928.

Doran, E., Fields, L., 2013. All stimuli are equal, but some are more equal than others: measuring relational preferences within an equivalence class. J. Exp. Anal. Behav. 98 (3), 243–256.

Doughty, A.H., Soydan, J.A., 2019. Differential derived stimulus relations across probe-trial versus adduction testing are not a function of comparison-stimulus presentation. Behav. Processes 166, 103903. https://doi.org/10.1016/j.beproc.2019.103903.

Dymond, S., May, R., 2018. Quantifying the empirical growth of relational frame theory research: a cautionary note. Psychol. Rec. 68 (2), 255–260. https://doi.org/10.1007/s40732-018-0278-z.

Fields, L., 2015. Stimulus relatedness in equivalence classes, perceptual categories, and semantic memory networks. Eur. J. Behav. Anal. 17 (1), 2–18. https://doi.org/10.1080/15021149.2015.1084713.

Fields, L., Moss, P., 2007. Stimulus relatedness in equivalence classes: interaction of nodality and contingency. Eur. J. Behav. Anal. 8 (2), 141–159. https://doi.org/10.1080/15021149.2007.11434279.

Fields, L., Landon-Jimenez, D., Buffington, D., Adams, B., 1995. Maintained nodal-distance effects in equivalence classes. J. Exp. Anal. Behav. 64 (2), 129–145. https://doi.org/10.1901/jeab.1995.64-129.

Fields, L., Arntzen, E., Nartey, R., Eilifsen, C., 2012. Effects of a meaningful, a discriminative, and a meaningless stimulus on equivalence class formation. J. Exp. Anal. Behav. 97 (2), 163–181. https://doi.org/10.1901/jeab.2012.97-163.

Finn, M., 2020. Exploring the Dynamics of Arbitrarily Applicable Relational Responding With the Implicit Relational Assessment Procedure [Dissertation, Ghent University]. http://hdl.handle.net/1854/LU-8654041.

Finn, M., Barnes-Holmes, D., McEnteggart, C., 2018. Exploring the Single-Trial-Type-Dominance-Effect in the IRAP: Developing a Differential Arbitrarily Applicable Relational Responding Effects (DAARRE) Model. Psychol. Rec. 68 (1), 11–25. https://doi.org/10.1007/s40732-017-0262-z.

Gavin, A., Roche, B., Ruiz, M., 2008. Competing contingencies over derived relational responding: a behavioral model of the Implicit Association Test. Psychol. Rec. 58 (3), 427–441. https://doi.org/10.1007/bf03395627.

Gawronski, B., De Houwer, J., 2014. Implicit measures in social and personality psychology. In: Reis, H., Judd, C. (Eds.), Handbook of Research Methods in Social and Personality Psychology, 2nd ed. Cambridge University Press, New York, pp. 283–310.

Greenwald, A., Banaji, M., 1995. Implicit social cognition: attitudes, self-esteem, and stereotypes. Psychol. Rev. 102 (1), 4–27. https://doi.org/10.1037//0033-295x.102.1.4.

Greenwald, A., Nosek, B., 2008. Attitudinal dissociation: what does it mean? In: Petty, R., Fazio, R., Brindol, P. (Eds.), Attitudes: Insights from the New Implicit Measures. Erlbaum, New Jersey, pp. 65–82.

Greenwald, A., McGhee, D., Schwartz, J., 1998. Measuring individual differences in implicit cognition: the implicit association test. J. Pers. Soc. Psychol. 74 (6), 1464–1480. https://doi.org/10.1037//0022-3514.74.6.1464.

Hughes, S., Barnes-Holmes, D., 2011. On the formation and persistence of implicit attitudes: new evidence from the implicit relational assessment procedure (IRAP). Psychol. Rec. 61 (3), 391–410. https://doi.org/10.1007/BF03395768.

Mizael, T., de Almeida, J., Silveira, C., de Rose, J., 2016. Changing racial Bias by transfer of functions in equivalence classes. Psychol. Rec. 66 (3), 451–462. https://doi.org/10.1007/s40732-016-0185-0.

Moss-Lourenco, P., Fields, L., 2011. Nodal structure and stimulus relatedness in equivalence classes: post-class formation preference tests. J. Exp. Anal. Behav. 95 (3), 343–368. https://doi.org/10.1901/jeab.2011.95-343.

Nicholson, E., Barnes-Holmes, D., 2012. The implicit relational assessment procedure (IRAP) as a measure of spider fear. Psychol. Rec. 62 (2), 263–277. https://doi.org/10.1007/bf03395801.

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. Psychol. Rec. 62 (3), 507–528. https://doi.org/10.1007/bf03395817.

O'Reilly, A., Roche, B., Gavin, A., Ruiz, M., Ryan, A., Campion, G., 2013. A function acquisition speed test for equivalence relations (FASTER). Psychol. Rec. 63 (4), 707–724. https://doi.org/10.11133/j.tpr.2013.63.4.001.

Parling, T., Cernvall, M., Stewart, I., Barnes-Holmes, D., Ghaderi, A., 2012. Using the implicit relational assessment procedure to compare implicit Pro-Thin/Anti-Fat attitudes of patients with anorexia nervosa and non-clinical controls. Eat. Disord. 20 (2), 127–143. https://doi.org/10.1080/10640266.2012.654056.

Power, P., Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., 2009. The implicit relational assessment procedure (IRAP) as a measure of implicit relative preferences: a first study. Psychol. Rec. 59 (4), 621–640. https://doi.org/10.1007/bf03395684.

Rehfeldt, R., Dymond, S., 2005. The effects of test order and nodal distance on the emergence and stability of derived discriminative stimulus functions. Psychol. Rec. 55 (2), 179–196. https://doi.org/10.1007/bf03395505.

Roche, B., O'Reilly, A., Gavin, A., Ruiz, M., Arancibia, G., 2012. Using behavior-analytic implicit tests to assess sexual interests among normal and sex-offender populations. Socioaffect. Neurosci. Psychol. 2 (1), 17335. https://doi.org/10.3402/snp.v2i0.17335.

Sidman, M., Kirk, B., Willson-Morris, M., 1985. Six-member stimulus classes generated by conditional-discrimination procedures. J. Exp. Anal. Behav. 43 (1), 21–42. https://doi.org/10.1901/jeab.1985.43-21.

Tyndall, I., Roche, B., James, J., 2004. The relation between stimulus function and equivalence class formation. J. Exp. Anal. Behav. 81 (3), 257–266. https://doi.org/10.1901/jeab.2004.81-257.

Urcuioli, P., 2013. Stimulus control and stimulus class formation. In: Madden, G., Dube, w., Hackenburg, T., Hanley, G., Lattal, K. (Eds.), APA Handbook of Behavior Analysis, Vol. 1. Methods and Principles. American Psychological Association, Washington, D.C., pp. 361–386.