BRIEF REPORT

DELAYED MATCHING-TO-SAMPLE TRAINING FACILITATES DERIVED RELATIONAL RESPONDING

Manish Vaidya and Kimberly N. Smith

UNIVERSITY OF NORTH TEXAS

Non-behavioral psychologists interested in the dynamics of learning make a distinction between procedures that enhance the acquisition of a set of skills and procedures that enhance the retention or transfer of that set of skills (Schmidt & Bjork, 1992; Schroth, 1995; 1997). The distinction is partly based on the observation that, under certain conditions, procedures that retard the acquisition of some performance during training can actually serve to enhance performance during transfer or retention tests (e.g., Schmidt, Young, Swinnen, & Shapiro, 1989; Shea & Morgan, 1979; see also Magill & Hall, 1990, for a review). For example, increasing the amount of task variability during training of a motor skill (patterned finger tapping) can retard acquisition but facilitate retention and transfer during tests (Hall & Magill, 1995). Likewise, decreasing the frequency of performance feedback during acquisition of a ballistic timing task impeded acquisition, but enhanced retention (Schmidt, et al., 1989). In general, this literature suggests that introducing certain task-relevant difficulties during the acquisition of a set of skills can serve to increase the likelihood of retention and effective transfer of those skills.

Research in stimulus equivalence provides a well-established preparation to study the dynamics reported above. First, the train and test preparation characteristic of research in stimulus equivalence provides the procedural context necessary to ask questions about the effects of learning conditions on the likelihood of transfer. Second, the definition of stimulus equivalence (Sidman & Tailby, 1982) provides a well-defined matrix of tasks which can be used to assess the extent and robustness of transfer and retention effects with less ambiguity relative to other procedures.

Arntzen (2006) investigated the effects of a set of imposed delays between the offset of sample stimuli and the onset of an array of comparison stimuli (hereafter, the retention interval) during training of baseline conditional relations on the likelihood of derived equivalence class formation with adult humans. Consistent with the literature cited above, Arntzen (2006) found a greater likelihood of equivalence-consistent responding as the duration of the retention interval was increased during training.

There were several features of Arntzen's (2006) procedure, however, that limit the conclusions that can be drawn from the results. For example, each participant was exposed to several conditions during which conditional relations were trained and emergent conditional relations assayed. Participants were exposed to either an ascending or descending series, in which the retention interval either increased or decreased, respectively, across conditions. Although the participants' performance in the ascending series improved with each increase in the retention interval value, the performance of participants' in the descending series was not similarly degraded as retention interval values were reduced. Arntzen suggests that the improvements seen in the ascending series may have been due, at least partially, to the increased exposure to training and testing circumstances rather than the imposed delays alone.

In the current study, we sought to minimize repeated training and testing opportunities by exposing an individual participant to only one

Address correspondence to Manish Vaidya at Department of Behavior Analysis, P.O. Box 310919, University of North Texas, Denton, TX. 76201, E-mail: vaidya@pacs.unt.edu.

delay and making the comparisons across groups of participants rather than within individual participants. In addition, we used common English and unfamiliar Portuguese and Czechoslovakian words as stimuli and tested only for the reversibility (symmetry) of trained conditional relations. To summarize, we asked whether introducing a delay between sample offset and comparison onset during the training of baseline conditional relations (e.g., given A1 pick B1, not B2, B3, or B4) would influence the likelihood of emergent symmetric relations (e.g., given B1 picking A1, not A2, A3, or A4).

METHOD

Participants

Twenty-six college students (10 male and 16 female) were recruited via publicly posted flyers and newspaper advertisements and were screened to ensure availability and a lack of familiarity with Behavior Analysis, and common words in Portuguese and Czechoslovakian because words from these languages comprised part of the set of experimental stimuli. Participants were randomly assigned to one of three groups (described below). Regardless of group membership, all participants earned $0.05 for every correct response during training and $1.00 for each completed session. The amounts earned were supplemented (at the end of the experiment when necessary) to maintain a minimum wage rate of $6.00 per hour.

Setting and Apparatus

Sessions were conducted in a 2 m by 3 m room equipped with a chair and a small table. The apparatus consisted of a Macintosh Power Book (G3) enclosed in a touch screen adapter (Troll Touch, Inc.). A custom-written program (MTS v. 11.6.7, see also Dube & Hirris, 1997) was used to present stimuli, manage all contingencies and collect data.

Procedure

The general plan of the experiment involved training a total of eight conditional relations – four with English words as sample stimuli and non-English words as comparison stimuli and the other four with non-English words as samples and English word comparisons – and testing for emergent reversibility of the trained conditional relations (see Table 1). The primary independent variable involved an experimenter-imposed delay between the offset of the sample stimulus and the onset of the comparison array during training and testing trials. Participants were divided into three groups: For Group 0 (n=8), Group 2 (n=8), and Group 8 (n = 10), there was a 0, 2 or 8 s delay, respectively, between the offset of the sample stimulus and the onset of the comparison array. The screen was white during the delay. All other contingencies were identical across groups and participants. All participants were instructed to respond by touching the screen. No other task-relevant instructions were presented.

Phase 1 (Baseline). Participants were exposed to one 8-trial block in which each conditional relation to be tested was presented once without programmed feedback. Trials began with the presentation of a stimulus in the vertical and horizontal center of the screen. Touching the sample stimulus removed it from the screen and, after passage of the programmed amount of time, produced an array of four comparison stimuli. During this phase, touching a comparison stimulus cleared the display and initiated a 1.5 s inter-trial interval (ITI). No other feedback was programmed.

Phase 2 (Training). Trials were identical to those presented in the baseline phase with the exception that choice of the experimenter-designated correct comparison stimuli resulted in the word "CORRECT" being displayed on the screen for 1 s accompanied by a tone and choice of any other comparison stimuli resulted in a 3 s time out during which the screen was white. Each of the eight conditional relations was randomly selected without replacement in a block of trials. Training trials were presented until the participant responded correctly on 24 consecutive trials or 384 training trials had been presented. In the latter case, the participant was invited to return for a second session which was identical to the first session.

Phase 3 (Testing). Participants who met the acquisition criterion were immediately exposed to testing trials in which stimuli previously functioning as comparison stimuli were presented as sample stimuli and vice versa. The delays between sample stimulus offset and comparison array onset remained at 0, 2, or 8 s depending upon group membership. The testing condition consisted of four exposures to 40-trial blocks (5 presentations of each of the 8 derived conditional relations). S1 was inadvertently exposed to only

Table 1.
The composition of training (left column) and testing (right column) trials during the experiment.

Training		Testing	
Sample	Comparison	Sample	Comparison
Food	Potrava	Potrava	Food
Bathroom	Koupelna	Koupelna	Bathroom
Ball	Esfera	Esfera	Ball
Farm	Fazenda	Fazenda	Farm
Postel	Bed	Bed	Postel
Duum	House	House	Duum
Comboio	Train	Train	Comboio
Margem	Bank	Bank	Margem

half of the testing trials relative to the remaining participants. Test trials were selected randomly (without replacement) such that each of the 8 test trial-types was presented once before any were repeated. Participation was considered complete after 160 testing trials or 2 hours, whichever came first.

Debriefing. Debriefing occurred in a separate session. Each participant completed a questionnaire regarding their use of verbal strategies as well an interview designed to get clarifications or follow-up answers. Upon completion, the participants were fully debriefed about the nature of the study.

RESULTS

Baseline. As a group, performance during the baseline block confirmed that participants were largely naïve with respect to the meaning of the Portuguese and Czechoslovakian words that served as experimental stimuli. The exceptions were S20 and S23 who selected the correct non-English words (presented as comparisons) on 3 of 4 baseline trials. Their performance on trials in which English words served as comparisons, however, was around chance levels of accuracy.

Acquisition. Participants in Groups 0, 2 and 8 met the training criterion of 24 consecutive correct responses in an average of 11.9 blocks (range = 5 to 24 blocks), 19.4 blocks (range = 9 to 61 blocks, and 11.4 blocks (range = 7 to 20 blocks), respectively. Participant 16 from Group 2 and Participants 24 and 25 from Group 8 failed to meet the training criterion by the end of their first 2 hour session and were invited back for a second session. Participants 24 and 25 from Group 8 failed to meet the training criterion during the second session as well and were paid and debriefed.

There was no difference in the rates at which conditional relations with English or non-English word samples were acquired within any of the three groups. There was also no significant difference in the speed of acquisition across groups (the average number of blocks to criterion in Group 2 are inflated due to one participant [S16] requiring 61 blocks to meet the acquisition criterion).

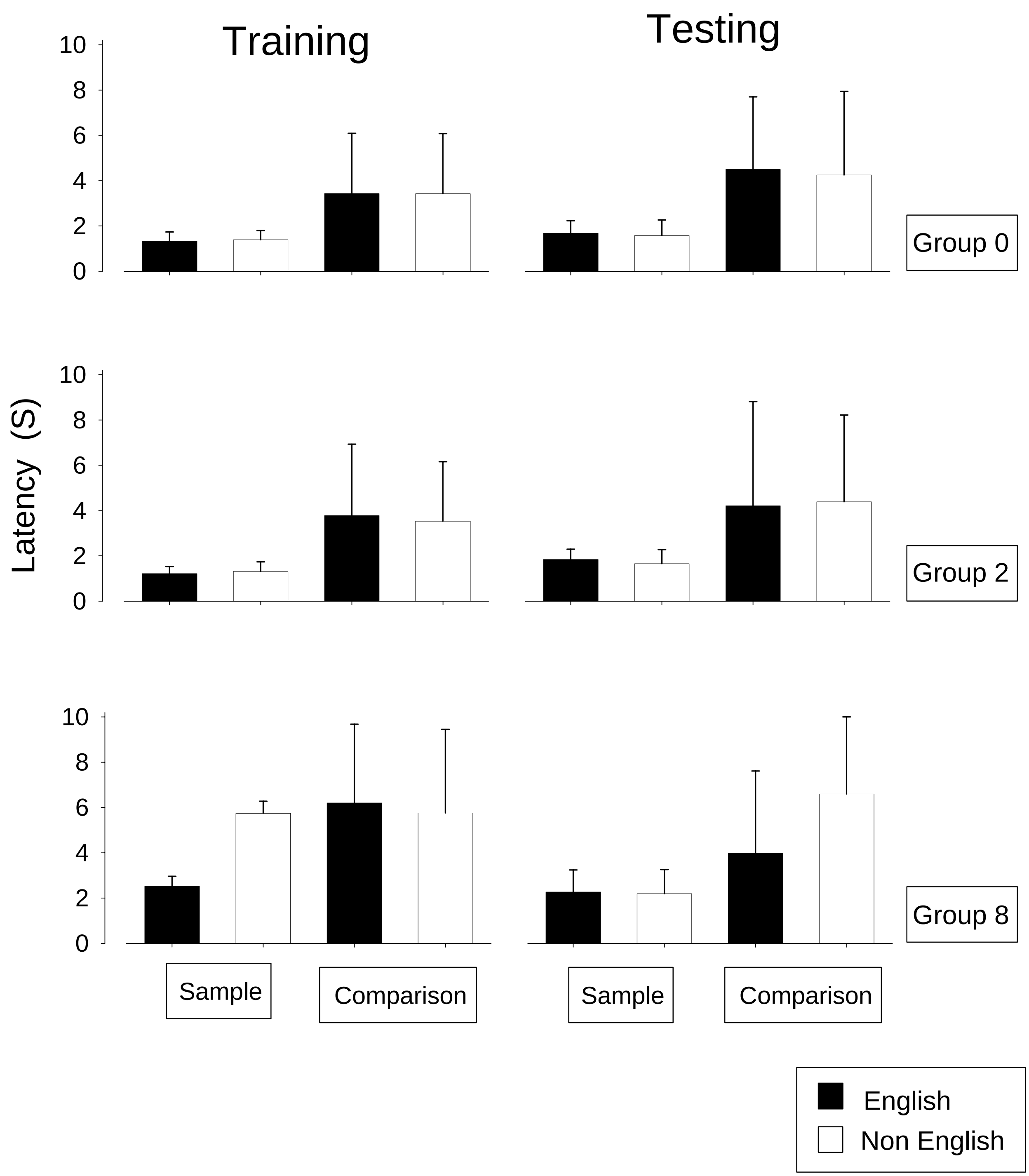

Figure 1

The left column in Figure 1 presents average observing-response and comparison–selection latencies during training separately for trials with English or non-English words serving as sample or comparison stimuli. Latencies for Groups 0, 2, and 8 are presented in the top, middle and bottom panels, respectively. The figure shows that, for Group 0 and Group 2, latencies to observe the sample stimulus or select a comparison stimulus were similar regardless of whether English or non-English words served as samples or comparisons. These data suggest that familiarity with the stimuli did not play a significant role in the organization of response latencies for participants in Group 0 and Group 2. The data from Group 8 differed, however. Participants in this group generally took longer to observe the sample stimulus and longer to make comparison selections. In addition, participants spent more time looking at non-English words (appearing as samples or comparisons) relative to participants in the other groups.

Figure 2 presents the percentage of correct trials across test blocks for each individual participant in Group 0 (left column), Group 2 (middle column), and Group 8 (right column). The participants' performances were deemed to show emergent symmetry if accuracy exceeded 80% (indicated by the solid horizontal reference line on each panel). The left column shows that accuracy on the test trials improved over the course of testing for three out of the eight participants (S3, S6, and S8) in Group 0. For four of the five remaining participants (S1, S2, S4, and S5), accuracy on the testing trials generally fluctuated around chance levels with no systematic difference in accuracy between trials involving English or foreign words as samples. S7's performance during tests was highly accurate in Block 2 but returned to chance levels of accuracy during Blocks 3 and 4.

The middle column shows that accuracy on the test trials improved over the course of testing for two participants (S10 and S13) and was highly accurate from the onset of testing for one participant (S16) in Group 2. The remaining five participants' performances (S9, S11, S12, S14, and S15) are more difficult to characterize briefly. S9 was highly accurate on trials with English word samples but not with non-English word samples. S11, S12 and S14's accuracy fluctuated around chance levels throughout the testing condition. S15's performance was moderately accurate (approximately 75% correct) during block 2 before returning to chance levels of accuracy.

The right column shows that performance on the test trials was highly accurate for five of the eight participants (S17, S18, S20, S22, and S23) in Group 8. In addition, S21's accuracy exceeded the established criterion for non-English samples during the final test block. Only two participants in this group (S19 and S26) failed to develop even one instance of symmetric relations among the stimuli. (An analysis of variance across groups, however, showed that the effect of delay was not significant, $F(2, 21) = 2.01$, $p > .05$). Interestingly, for all participants in this group, accuracy on trials involving non-English word samples was greater than accuracy on trials involving English word samples. In addition, for four of the eight participants (S17, S18, S20, and S22), performance on trials involving non-English word samples (and English word comparisons) was at or near 100% accurate from the beginning and throughout the testing condition. This may have been due to participants' familiarity with the commonly used English words comprising the comparison array. Alternatively, this performance may reflect differences in learning during the training condition when English words served as sample stimuli and the non-English words served as comparisons.

The right column of Figure 1 presents average latencies to respond to the sample stimulus and to select comparison stimuli for each of the three groups during testing. Mean latencies for Group 0, Group 2, and Group 8 are presented in the top, middle, and bottom panels, respectively. The data presented in the top and middle panels show that, for participants in Group 0 and Group 2, there was no systematic difference in the latencies to respond to English or non-English words presented as sample or comparison stimuli. The participants in Group 8, however, generally took longer to respond to all stimuli. In particular, participants in this group took longer to respond to non-English word comparisons relative to English word comparisons during testing.

A large majority of the participants reported the use of verbal devices during the training condition. These included memorization or rehearsal (13 participants) and intraverbal naming (7 participants). Examples of intraverbal naming included phrases such as "merging banks" or "pots cook food". There was, however, no obvious relation between the kind of naming strategy employed and the likelihood of emergent symmetrical relations. For example, of the seven participants that reported some intraverbal naming, three showed emergent symmetry (S8, S18, and S20) and four did not (S1, S4, S5, S11).

DISCUSSION

The purpose of the present study was to evaluate the effects of delays between sample offset and comparison array onset on the acquisition of conditional discriminations and the

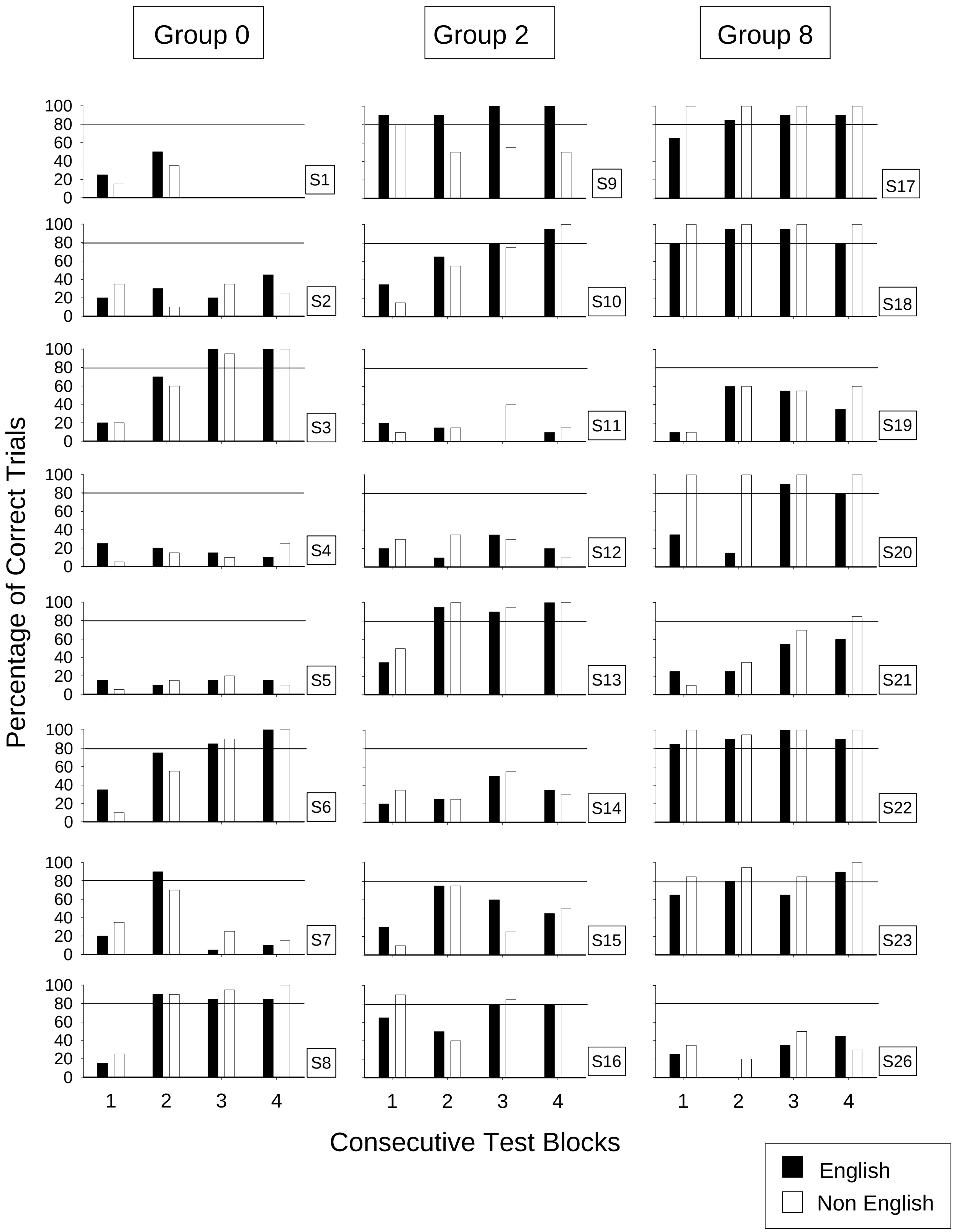
Group 0
Group 2
Group 8
Percentage of Correct Trials
Consecutive Test Blocks
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
S12
S13
S14
S15
S16
S17
S18
S19
S20
S21
S22
S23
S26
English
Non English

Figure 2

development of their symmetrical counterparts. The data show some retardation in acquisition – participants either required more training (S16 in Group 2) or failed to acquire the conditional relations entirely (S24 and S25 in Group 8) as the delay between sample offset and comparison onset was increased. For the most part, however, there was no difference in the number of trials required to meet criterion across groups. These data are inconsistent with the behavioral literature on delay matching-to-sample (Sagrisson & White, 2001) as well as the non-behavioral literature cited above and may provide an interesting follow-up question for future research. Perhaps the delays imposed between sample offset and comparison array onset were too short to produce a decrement in learning rates.

The data also show that participants exposed to longer delays between sample offset and comparison onset during training are more likely to make symmetry-consistent choices on test trials. These data provide a systematic replication of the results reported by Arntzen (2006) and are also consistent with other data suggesting that the introduction of task-relevant difficulties during training can facilitate retention and transfer of learned skills (Magill & Hall, 1990).

One possible factor that may have influenced the differential development of emergent relations is the duration of contact with the experimental stimuli. The data show that participants in Group 8 spent nearly twice as long in the presence of the sample and comparison stimuli relative to participants in Groups 0 and 2. The extended exposure to sample and comparison stimuli may have served to eliminate irrelevant sources of control and increased control by experimenter-designated features such that symmetrical relations became more likely. This interpretation is consistent with the predictions of Stimulus Control Topography Coherence (hereafter, SCTC) theory (McIlvane & Dube, 2003). According to SCTC theory, emergent conditional relations require coherence of experimenter-defined and obtained stimulus control topographies. With respect to the current data set, long delays between sample offset and comparison onset during training may have served to increase the time participants spent observing stimuli prior to a sample-observing or comparison-selection response. The additional interaction with the stimuli may have facilitated the development of stimulus control topography coherence and led to an increased likelihood of emergent symmetry-consistent choices on the test trials.

An alternative or complementary mechanism may have been the development of other pre-current or supplemental behavior (such as naming) which facilitated the emergence of symmetrical relations (cf. Horne & Lowe, 1996). The role of naming in the current data set remains unclear, however. Although 33% of the participants reported some intraverbal naming during the training conditions, only three of those seven participants made symmetry-consistent choices on the test trials. Future research should attempt to manipulate the likelihood of naming directly to elucidate its role in the organization of conditional discrimination performances.

In conclusion, the methodology of equivalence research appears to provide a sensitive preparation to investigate the relation between training and transfer reported by researchers outside the behavioral tradition. These procedures also have the potential to more broadly inform our understanding of the derivation of relational responding. A systematic database relating the circumstances of training to the likelihood of derived relational responding may help to identify the conditions that are necessary and sufficient for such responding to emerge. In addition to the important theoretical benefits, such work may also contribute to the development of technologies designed to facilitate derived relational responding – an important outcome in a large variety of clinical and educational settings.

REFERENCES

Arntzen, E. (2006). Delayed matching to sample: Probability of stimulus equivalence as a function of delays between sample and comparison stimuli during training. *The Psychological Record,* **56**, 135-167.

Dube, W., & Hirris, J. (1997). Match to Sample Program (Version 11.6.7) [Computer software]. Waltham, MA: E. K. Shriver Center for Mental Retardation.

Hall, K.G. & Magill, R.A. (1995). Variability of practice and contextual interference in motor skill learning. *Journal of Motor Behavior*. **27**, 299-309.

Horne, P. J. & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal* *of the Experimental Analysis of Behavior,* **65**, 185-241.

Magill, R.A. & Hall, K.G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science,* **9**, 241-289.

McIlvane, W. J. & Dube, W. V. (2003). Stimulus control topography coherence theory: Foundations and extensions. *The Behavior Analyst,* **26**, 195-213.

Sargisson, R. J., & White, K. G. (2001). Generalization of delayed matching to sample following training at different delays. *Journal of the Experimental Analysis of Behavior,* **75**, 1-14.

Schmidt, R.A & Bjork, R.A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science,* **3**, 207-217.

Schmidt, R.A., Young, D.E., Swinnen, S., & Shapiro, D.C. (1989). Summary knowledge of results for skill acquisition: Support for the guidance hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15**, 352-359.

Schroth, M. L. (1995). Variable delay of feedback procedures and subsequent concept formation transfer. *The Journal of General Psychology,* **122**, 393-399.

Schroth, M. L. (1997). The effects of different training conditions on transfer in concept formation. *The Journal of General Psychology,* **124**, 157-165.

Shea, J.B. & Morgan, R.L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory,* **5**, 179-187.

Sidman, M. & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior,* **37**, 5-22.