

A PRECURSOR TO THE RELATIONAL EVALUATION PROCEDURE: ANALYZING STIMULUS EQUIVALENCE

VERONICA A. CULLINAN and DERMOT BARNES
University College Cork, Ireland

PAUL M. SMEETS
Leiden University, The Netherlands

Four experiments are described that employed a form of go/no go procedure that we refer to as a precursor to the Relational Evaluation Procedure (pREP). The pREP was used to train and test for equivalence responding, and the results of this procedure were compared with the results achieved using a standard matching-to-sample procedure. Each trial in the pREP involved a sample stimulus and either a positive or negative comparison stimulus being presented successively on a computer screen, followed by a 5-sec response interval. Subjects were required to press the space bar of a computer keyboard on sample-positive comparison trials and to not press the space bar on sample-negative comparison trials. Using *either* the pREP or a matching-to-sample procedure 20 subjects were trained in the following four tasks, $A1 \rightarrow B1$, $A2 \rightarrow B2$, $B1 \rightarrow C1$, $B2 \rightarrow C2$. They were then tested for four symmetrical relations ($B1 \rightarrow A1$, $B2 \rightarrow A2$, $C1 \rightarrow B1$, $C2 \rightarrow B2$), and two equivalence relations ($C1 \rightarrow A1$ and $C2 \rightarrow A2$), using *both* the pREP and matching-to-sample procedures, with the order of presentation of the two types of test varied across experiments. The results of the four experiments reported here demonstrate that the pREP is less effective than the matching-to-sample procedure in generating equivalence responding. However, performance on the pREP tests improved when the subjects had prior exposure to matching-to-sample training and/or testing. This finding suggests that further study of the interactions between these two procedures, combined with suggested refinements to the pREP itself, may contribute to a fuller understanding of those variables most relevant to producing equivalence responding.

Researchers studying the phenomenon of stimulus equivalence often use a matching-to-sample (MTS) procedure. In a typical experiment, for example, four MTS tasks may be used during the training phase to establish related conditional discriminations. On each of these

This research was conducted as part of Veronica Cullinan's doctoral research program under the supervision of Dermot Barnes. Requests for reprints should be addressed to Dermot Barnes, Behavior Analysis and Cognitive Science Unit, Department of Applied Psychology, University College Cork, Cork, Ireland (e-mail: Lemmy@ucc.ie).

tasks, a single sample stimulus is presented along with two comparison stimuli (alphanumeric labels are often used to designate these stimuli), and the subject is required to choose one of the comparisons. During the training, choosing stimulus B1 when stimulus A1 is presented as a sample produces positive feedback (e.g., the word "Correct" is presented). Choosing stimulus B2, however, produces negative feedback (e.g., the word "Wrong"). In this way, the following four MTS performances may be established: $A1 \rightarrow B1$, $A2 \rightarrow B2$, $B1 \rightarrow C1$, $B2 \rightarrow C2$. After a subject has reached a predetermined mastery criterion (e.g., 12 consecutively correct responses), he or she is exposed to an equivalence test. During this test, six MTS tasks may be used to probe (i.e., without feedback) for the emergence of responding in accordance with symmetry and combined symmetry and transitivity. Symmetry requires that the conditional discriminations be functionally reversible (i.e., $B1 \rightarrow A1$, $B2 \rightarrow A2$, $C1 \rightarrow B1$, $C2 \rightarrow B2$). Transitivity requires that the $A \rightarrow B$ and $B \rightarrow C$ conditional discriminations produce $A \rightarrow C$ responding (i.e., $A1 \rightarrow C1$, $A2 \rightarrow C2$), and combined symmetry and transitivity requires that the $A \rightarrow B$ and $B \rightarrow C$ conditional discriminations produce $C \rightarrow A$ responding (i.e., $C1 \rightarrow A1$ and $C2 \rightarrow A2$). When these untrained matching performances emerge, they are often described as equivalence responding.

Recently many studies on emergent conditional discrimination performances have used alternatives to the standard MTS procedure. For example, complex stimuli have been used in identity and arbitrary MTS (see Stromer, McIlvane, & Serna, 1993; Markham & Dougher, 1993; Schenk, 1993; Smeets, Schenk, & Barnes, 1995; Smeets & Streifel, 1994). A complex stimulus is made up of more than one element; for example, a color superimposed on a form. Using complex stimuli as either samples or comparisons in MTS training, allows for testing of the relations between the elements of each complex stimulus as well as between the complex stimuli and other single element stimuli. For example, a complex sample and single element comparison could be used to train the relation $AB \rightarrow C$ (e.g., color-form sample and form only comparison). Some of the relations which might be generated by this training are $A \rightarrow B$, $A \rightarrow C$, and $B \rightarrow C$. In general, these studies have shown that matching to complex stimuli generates conditional relations between the elements of the complex stimulus, as well as between those complex stimuli and other single element stimuli. In some cases the emergent relations observed using complex stimuli have been characterized as equivalence relations (but see, Stromer et al., 1993).

Another alternative to the standard MTS procedure was developed by Leader, Barnes, and Smeets (1996). The basic procedure involved stimulus pairs being presented sequentially with varying inter-pair and intra-pair temporal gaps. Subjects were not required to demonstrate any overt response, but were instructed simply to look at the computer screen on which the stimuli were presented. After a predetermined number of exposures to the training procedure, subjects were presented with a

standard MTS equivalence test. In a series of experiments adult subjects consistently demonstrated equivalence responding using this procedure.

The Leader et al. (1996) training procedure proved to be remarkably effective in generating emergent matching performances. However, because their training procedure did not require the subject to emit an overt response, it would be difficult to use it to *test* for equivalence responding. In an attempt to extend the work conducted by Leader et al., the present study developed a procedure for examining equivalence class formation that employed neither MTS training *nor* testing. Specifically, a precursor to the Relational Evaluation Procedure (pREP)¹ was used to present a series of conditional discrimination tasks. In the pREP each trial consists of the presentation of one sample followed by either a positive comparison or a negative comparison, with a go/no-go response requirement on each trial (see D'Amato & Colombo, 1985). In effect, subjects are trained to press the space bar of a computer keyboard on 'correct' trials (i.e., when the sample is followed by the predesignated positive comparison) and to not press the space bar on 'incorrect' trials (i.e., when the sample is followed by a negative comparison). Pressing the space bar can be seen as analogous to selecting a comparison in a MTS procedure and not pressing as analogous to not selecting a comparison in a MTS procedure (i.e., positive and negative sample-comparison relations are demonstrated on separate trials). The four experiments reported here compare training and testing using the pREP, with training and testing using a standard MTS procedure. Apart from developing another procedure for producing emergent matching performances, comparing the relative effectiveness of these two procedures may shed some light on those variables responsible for emergent matching behaviors.

General Method

Subjects

Twenty students, 15 female and 5 male, attending University College Cork, participated in the experiments. Their ages ranged from 18 to 30 years, and they had no prior experience with stimulus equivalence research. They were recruited through notice-board advertisements and personal contacts, and they were randomly assigned to one of the four experiments (i.e., 5 subjects in each experiment).

Apparatus and Setting

All subjects were trained and tested individually in a quiet room free of distractions. The stimuli used were the nonsense syllables ZID, JOM, ROG, CUG, BEH, DAX which will be represented here by the alphanumerics A1, B1, C1, A2, B2, C2. Stimuli were presented on an

¹The reader is referred to Hayes and Barnes (1997) for a detailed treatment of the Relational Evaluation Procedure itself, and in particular the methodological and conceptual concerns underlying its development.

Apple Macintosh LC III Computer, and subjects responded by pressing various marked keys on the keyboard. The computer was programmed in BBC BASIC to control the presentation of stimuli and to record responses. Training and testing trials were presented in blocks, and after each block of trials subjects were free to move around or take a break, if

EXPERIMENTAL PROCEDURE

pREP TRAINING AND TESTING

Train

1 sec	A1	A2	B1	B2	A1	A2	B1	B2
1 sec								
1 sec	B1	B2	C1	C2	B2	B1	C2	C1
5 sec	press	press	press	press	no press	no press	no press	no press

Test Sym

1 sec	B1	B2	C1	C2	B2	B1	C2	C1
1 sec								
1 sec	A1	A2	B1	B2	A1	A2	B1	B2
5 sec	?	?	?	?	?	?	?	?
	press	press	press	press	no press	no press	no press	no press

Test Equiv

1 sec	C1	C2	C1	C2
1 sec				
1 sec	A1	A2	A2	A1
5 sec	?	?	?	?
	press	press	no press	no press

MTS TRAINING AND TESTING

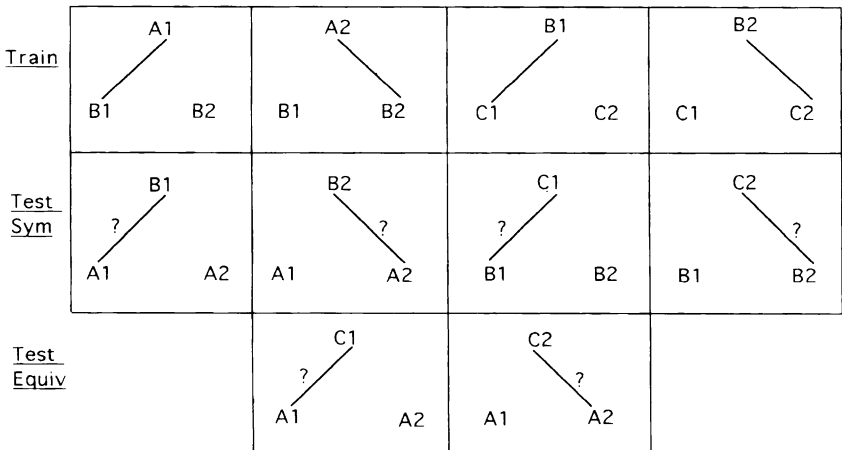


Figure 1. Schematic representation of training and testing tasks used in the MTS procedure and the pREP.

they wished, while the experimenter checked the results. Sessions ranged from 30 min to 3 hr, and almost all subjects required more than one session. (This wide variation in duration of experimental sessions was caused by the fact that the overall procedures were quite time consuming for most subjects, and thus flexibility was required to fit experimental sessions into their academic timetables.)

Procedure

General experimental sequence. The overall experimental design involved each subject being trained in a series of conditional discriminations using *either* a standard MTS procedure, *or* the pREP. All subjects were trained in the following relations: $A1 \rightarrow B1$, $A2 \rightarrow B2$, $B1 \rightarrow C1$, $B2 \rightarrow C2$, irrespective of the training procedure to which they were exposed. When a subject reached the training criterion, he or she was tested for the emergence of symmetrical ($B1 \rightarrow A1$, $B2 \rightarrow A2$, $C1 \rightarrow B1$, $C2 \rightarrow B2$) and equivalence ($C1 \rightarrow A1$, $C2 \rightarrow A2$) responding, with both trial types (i.e., symmetry and equivalence) mixed in each block of test trials. Pilot work had shown that the experiment would require a large number of exposures to the training and testing procedures, and thus it was decided not to test for both transitivity and equivalence in order to keep the time required to a minimum. All subjects were presented with *both* standard MTS tests *and* pREP tests. Baseline conditional discriminations were retrained (to criterion) before *each* test in each of the four experiments. The order of presentation of the training and testing procedures will be described separately for each experiment.

Precursor to the Relational Evaluation Procedure. For the pREP subjects were seated in front of the computer monitor and the instructions were read aloud to them. During training phases the instructions were as follows:

One nonsense syllable will appear in the centre of the screen for one second, the screen will clear for one second, then another nonsense syllable will appear for one second. Then there will be a five second pause. During this five second pause I want you to either press the space bar, or not press the space bar. A message will appear on the screen saying either "GOOD" and adding a point on to a running total which you will see on the screen, or "BAD" and subtracting a point from the running total. When the experiment is finished you will be paid a penny for each point you earn, so you should try to earn as many points as possible. When this session is finished a message will appear asking you to call the experimenter. I will be waiting outside. Do you have any questions?

During testing phases the instructions were:

In this part of the experiment one nonsense syllable will appear in the centre of the screen for one second, the screen will clear for

one second, then another nonsense syllable will appear for one second, as before. Then there will be a five second pause. During this five second pause I want you to either press the space bar, or not press the space bar. However this time you will not get a message saying good or bad, so just do whatever you think is right. Do you have any questions?

Any questions asked were answered by repeating the relevant section of the instructions, and then the experimenter left the room.

In both training and testing phases the sample stimulus appeared in the center of the screen for 1 s, the screen then cleared for 1 s, and either a positive or negative comparison stimulus was presented for 1 s. There was then a 5-s response interval during which the subject was required either to press the space bar or not press the space bar. If the subject pressed the space bar the response interval was immediately terminated and either the programmed consequences were presented (in training phases) or the next trial was presented (in test phases). Training trials were presented in blocks of 40 trials with each of the 8 tasks (A1 → B1, A1 → B2, A2 → B1, A2 → B2, B1 → C1, B1 → C2, B2 → C1, B2 → C2) presented five times in a quasi-random order. Test trials were presented in blocks of 120 trials, with each of the 12 tasks (B1 → A1, B1 → A2, B2 → A1, B2 → A2, C1 → B1, C1 → B2, C2 → B1, C2 → B2 [symmetry]; C1 → A1, C1 → A2, C2 → A1, C2 → A2 [equivalence]) being presented 10 times in quasi-random order (see Figure 1).

Matching-to-sample procedure. During exposure to the MTS procedure the experimental setting was identical to that for the pREP. For training phases the instructions were as follows:

Some nonsense syllables will appear on the screen, one at the top, and two at the bottom. I want you to look at the nonsense syllable at the top and then choose one of the nonsense syllables at the bottom. To choose the one on the left press the marked key on the left, to choose the one on the right press the marked key on the right. When you have made your choice a message will appear on the screen saying either "GOOD" and adding a point on to a running total which you will see on the screen, or "BAD" and subtracting a point from the running total. When the experiment is finished you will be paid a penny for each point you earn, so you should try to earn as many points as possible. At the end of this session a message will appear asking you to call the experimenter, I will be waiting outside. Do you have any questions?

During testing phases the instructions were:

In this part of the experiment one nonsense syllable will appear at the top of the screen and two at the bottom of the screen, as before. I want you to look at the nonsense syllable at the top and choose one of the ones at the bottom again. However this time you will not get a message saying good or bad, so just do whatever you think is right. Do you have any questions?

Any questions asked were answered by repeating the relevant section of the instructions, and then the experimenter left the room.

In both training and testing phases the sample stimulus was presented near the top of the screen, followed 1 s later by two comparison stimuli presented to the left and right of the sample along the lower edge of the screen. All three stimuli remained visible on the screen until the subject made a response. Subjects responded by pressing one of two marked keys on the keyboard (the letters 'Z' and 'M'). During training phases a response was followed by presentation of programmed consequences; during testing there were no programmed consequences. Training trials were presented in blocks of 40 trials with each of the 4 tasks (A1 → B1/B2, A2 → B1/B2, B1 → C1/C2, B2 → C1/C2) presented 10 times in a quasi-random order, with the position of comparison stimuli counterbalanced across trials. Test trials were presented in blocks of 60 trials with each of the 6 tasks (B1 → A1/A2, B2 → A1/A2, C1 → B1/B2, C2 → B1/B2 [symmetry], C1 → A1/A2, C2 → A1/A2 [equivalence]) being presented 10 times in quasi-random order, again with the position of comparisons counterbalanced across trials (see Figure 1).

Programmed consequences. When training subjects using either of the two procedures, feedback was provided on the screen following all responses (i.e., choosing the left or right comparison during MTS, and pressing or not pressing the space-bar during the pREP). If a response was designated as correct then an auditory tone was presented and "GOOD: POINTS = XX" appeared on the screen, displaying a total of points earned during that training block, incremented by one. If the response was designated as incorrect then an auditory tone was presented and "BAD: POINTS = XX" appeared on the screen, displaying the total points earned during that training block decremented by one. Upon completion of the entire study, subjects were paid a penny for each point earned. During testing phases there were no programmed consequences. A 2-s intertrial interval (blank screen) separated all trials.

Training, test, and stability criteria. For both MTS and the pREP the training criterion was 90% correct, that is 36 trials correct out of 40 in any one block of training trials. Once this criterion was met subjects were presented with a test phase. However, if time constraints prevented a subject being exposed to the test in the same session, he or she was retrained to criterion at the beginning of the next session, immediately prior to being presented with a test.

Each test task was presented 10 times (in a quasi-random order) in each test block. The criterion for passing a test was 8 out of 10 correct responses on each testing task, in any one block of test trials. This criterion amounts to 80% or more correct. Pilot work had shown that due to the large number of test trials involved, especially in the pREP, subjects often reported making 'slip of the hand' errors because of fatigue and loss of concentration, and thus an 80% criterion was deemed to be more appropriate than the 90% used for training. If this test criterion was not reached, the subject was retrained to criterion and then presented with the same test again. This continued until the subject

either passed the test or demonstrated a stable incorrect performance on two consecutive test presentations. The criterion for a stable, incorrect performance required that the difference between scores on each individual test task, across the 2 blocks of test trials, be no more than 2 (out of a possible total of 10). After either passing the test or demonstrating a stable incorrect performance, the subject was again retrained and then presented with the alternative test, and if necessary, retrained and retested until he or she passed or demonstrated a stable, incorrect performance. Finally, the subject was reexposed to the test that had first been presented (see Figure 2 for a schematic representation of the training and testing sequence employed across all four experiments).

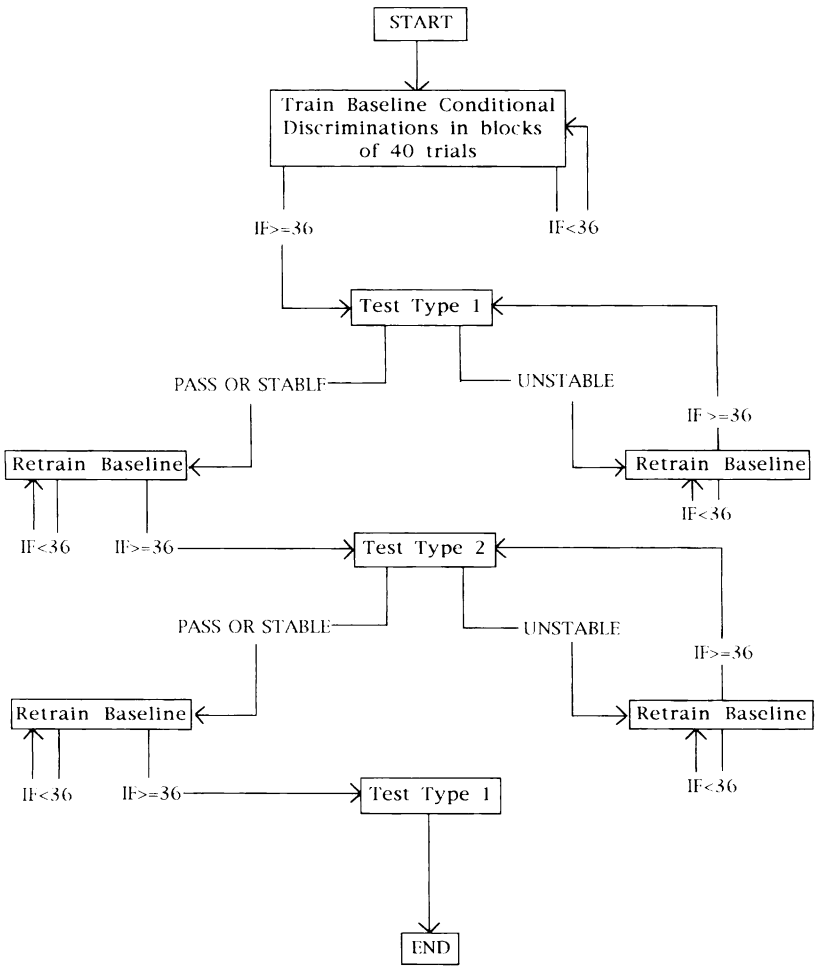


Figure 2. Schematic representation of the training and testing sequence employed across the four experiments (see text for details).

In general, subjects found the procedures described here quite difficult and time consuming, and for this reason there were occasional minor deviations from the planned experimental sequences. For example, on a few occasions subjects received only 16 retraining trials instead of 40 in between exposures to tests. This occurred only in cases where a subject had consistently completed previous 40-trial retraining blocks without error, and in all these cases subjects completed the 16 retraining trials without error. Similarly, on one occasion (Experiment 2, Subject 3) a subject received only one exposure to the pREP test because this subject could not guarantee future availability for participation in the study, and therefore the experimenter decided to ignore the stable incorrect criterion in this case and allow the subject to complete the experiment.

Experiment 1

Procedure

All five subjects were trained to criterion using the pREP. They were trained to press the space bar when presented with the four tasks $A1 \rightarrow B1$, $A2 \rightarrow B2$, $B1 \rightarrow C1$, $B2 \rightarrow C2$, and to not press the space bar when presented with the four tasks $A1 \rightarrow B2$, $A2 \rightarrow B1$, $B1 \rightarrow C2$, $B2 \rightarrow C1$ (see Figure 1). When subjects had reached the training criterion of 90% correct responding on a block of trials, they were presented with the pREP test. This consisted of eight symmetry tasks and four equivalence tasks. Four of the symmetry tasks (i.e., $B1 \rightarrow A1$, $B2 \rightarrow A2$, $C1 \rightarrow B1$, $C2 \rightarrow B2$) required subjects to respond by pressing the space bar. The other four tasks (i.e., $B1 \rightarrow A2$, $B2 \rightarrow A1$, $C1 \rightarrow B2$, $C2 \rightarrow B1$) required subjects not to press the space bar during the response interval. Similarly, two of the equivalence tasks (i.e., $C1 \rightarrow A1$, $C2 \rightarrow A2$) required subjects to press the space-bar, and the other two equivalence tasks (i.e., $C1 \rightarrow A2$, $C2 \rightarrow A1$) required subjects to not press the space bar during the response interval. If subjects did not reach the test criterion of 8 out of 10 correct responses on each test task, they were retrained using the pREP and then reexposed to the pREP test. This retraining and retesting continued until subjects either reached the test criterion or demonstrated a stable, incorrect performance (as defined in the General Procedure). Subjects were then trained again using the pREP and were presented with a MTS test. This tested for the same relations as the pREP test, but on each trial two comparisons were presented with each sample (e.g., in a symmetry trial B1 might be presented as a sample with A1 and A2 as comparisons). Therefore, the MTS test consisted of four symmetry tasks and two equivalence tasks (see Figure 1). If subjects did not reach the test criterion on this MTS test they were again retrained (using the pREP) and retested until they either reached the test criterion or demonstrated a stable incorrect performance. Finally, subjects were retrained to criterion and tested using the pREP.

Results and Discussion

Table 1

Initial Training Trials, Number of Exposures to Tests, and Results for Symmetry and Equivalence on Final Exposure to Each Type of Test

Experiment 1		Subjects				
		1	2	3	4	5
pREP Train	Tr. trials to Criterion	80	520	560	120	640
pREP Test	No. of Exposures to Test	4	6	3	1	2
	Results of Final Exp.-Sym.	PASS	PASS	PASS	PASS	PASS
MTS Test	-Equiv.	FAIL	FAIL	FAIL	PASS	FAIL
	No. of Exposures to Test	6	3	2	1	1
pREP Test	Results of Final Exp.-Sym.	PASS	PASS	PASS	PASS	PASS
	-Equiv.	PASS	FAIL	PASS	PASS	PASS
pREP Test	No. of Exposures to Test	1	1	1	1	1
	Results of Final Exp.-Sym.	PASS	PASS	PASS	PASS	PASS
pREP Test	-Equiv.	FAIL	FAIL	FAIL	PASS	FAIL
Experiment 2		Subjects				
		1	2	3	4	5
pREP Train	Tr. trials to Criterion	440	80	480	200	200
MTS Test	No. of Exposures to Test	5	1	8	5	6
	Results of Final Exp.-Sym.	PASS	PASS	PASS	PASS	PASS
pREP Test	-Equiv.	FAIL	PASS	FAIL	PASS	PASS
	No. of Exposures to Test	3	5	1	3	3
pREP Test	Results of Final Exp.-Sym.	PASS	PASS	PASS	PASS	PASS
	-Equiv.	FAIL	PASS	FAIL	PASS	FAIL
MTS Test	No. of Exposures to Test	1	1	1	1	1
	Results of Final Exp.-Sym.	PASS	PASS	PASS	PASS	PASS
MTS Test	-Equiv.	FAIL	PASS	FAIL	PASS	PASS
Experiment 3		Subjects				
		1	2	3	4	5
MTS Train	Tr. trials to Criterion	80	120	120	120	80
pREP Test	No. of Exposures to Test	2	1	3	2	7
	Results of Final Exp.-Sym.	PASS	PASS	PASS	PASS	PASS
MTS Test	-Equiv.	PASS	PASS	PASS	PASS	FAIL
	No. of Exposures to Test	2	1	1	1	1
pREP Test	Results of Final Exp.-Sym.	PASS	PASS	PASS	PASS	PASS
	-Equiv.	PASS	PASS	PASS	PASS	PASS
pREP Test	No. of Exposures to Test	1	1	1	1	1
	Results of Final Exp.-Sym.	PASS	PASS	PASS	PASS	PASS
pREP Test	-Equiv.	PASS	PASS	PASS	PASS	FAIL
Experiment 4		Subjects				
		1	2	3	4	5
MTS Train	Tr. trials of Criterion	440	160	160	280	40
MTS Test	No. of Exposures to Test	1	1	1	4	1
	Results of Final Exp.-Sym.	PASS	PASS	PASS	FAIL	PASS
pREP Test	-Equiv.	PASS	PASS	PASS	FAIL	PASS
	No. of Exposures to Test	1	1	2	3	2
pREP Test	Results of Final Exp.-Sym.	PASS	PASS	PASS	PASS	PASS
	-Equiv.	PASS	PASS	FAIL	FAIL	PASS
MTS Test	No. of Exposures to Test	1	1	1	1	1
	Results of Final Exp.-Sym.	PASS	PASS	PASS	FAIL	PASS
MTS Test	-Equiv.	PASS	PASS	PASS	FAIL	PASS

Table 1 shows the individual subject data for the initial pREP training, the total number of tests to which each subject was exposed, and the results of the final exposure to each type of test (a detailed breakdown of each subject's performance is shown in Appendix A). For the purposes of communication, a subject who reached the test criterion is referred to here as having 'passed' that test. A stable incorrect performance is referred to as a 'fail.' The number of pREP training trials required to reach criterion ranged from 80 (S1) to 640 (S5) trials. All subjects responded in accordance with the symmetry relations on the pREP test (on the first or second exposure). However, even after up to six exposures to the test, 4 of the 5 subjects failed to respond in accordance with the equivalence relations on the pREP test. The one subject who passed the entire test (S4) did so on the first exposure. When subjects were then presented with a MTS test, all 5 subjects again produced symmetry responding, but on this test 4 of the 5 subjects also responded in accordance with the equivalence relations. Subject 1 required six exposures to the MTS test to achieve this, the other subjects who passed required only one or two exposures. When they were then reexposed to the pREP test, the results were similar to the original pREP tests, in that all subjects produced symmetry responding, and only 1 subject (S4) of 5 produced equivalence.

Overall the pREP training and testing proved to be very effective in producing symmetry responding, but relatively ineffective in producing equivalence responding. Furthermore, the MTS test appeared to produce equivalence more readily than the pREP. Because the MTS test apparently facilitated equivalence responding, it was decided to repeat Experiment 1 but to present the MTS test before the pREP test. Would earlier exposure to the MTS test (relative to Experiment 1) facilitate equivalence responding on a subsequent pREP test?

Experiment 2

Procedure

All training and testing procedures and criteria used in Experiment 2 were identical to Experiment 1, but the order in which the two types of tests were presented was different. Subjects were trained to criterion using the pREP training. Then they were exposed to a test using a standard MTS test procedure. If they did not pass the test on the first exposure, they were retrained and retested until they passed or demonstrated a stable, incorrect performance (as defined in the General Procedure). They were then retrained and exposed to the pREP test. If they did not pass this test on the first exposure they were retrained and retested until they passed or demonstrated a stable, incorrect performance. Finally they were exposed to the MTS test again.

Results and Discussion

Table 1 shows individual subject data for Experiment 2 (see

Appendix B for a detailed breakdown of each subject's performance). Subjects required between 80 (S2) and 480 (S3) training trials to reach the training criterion. On the MTS test all subjects responded in accordance with symmetry relations, and 3 of the 5 subjects produced equivalence responding. Subject 2 passed both symmetry and equivalence on the first exposure to the test, Subjects 4 and 5 passed the entire test on the fifth and sixth exposures, respectively. The two subjects who failed to produce equivalence (S1 and S3) required five and eight exposures to the test, respectively, before demonstrating a stable, incorrect performance. On the pREP test, all 5 subjects produced symmetry, and 2 subjects produced equivalence. Subject 5, having produced equivalence responding on the earlier MTS test, failed to demonstrate equivalence when presented with the pREP test. When reexposed to the MTS test, the subjects' performances reverted back to that demonstrated on the initial MTS test, with all subjects producing symmetry, and three subjects producing equivalence responding.

These results demonstrate that even when subjects are presented with a MTS test immediately after pREP training to criterion, equivalence responding does not readily emerge, although, as in Experiment 1, symmetry responding was universal. Nevertheless, the fact that 3 subjects produced equivalence on the MTS test (in contrast to 2 on the pREP test) suggests once more that MTS may be more effective than the pREP for producing equivalence. To explore this issue further, the next experiment replicated Experiment 1, but subjects were trained using MTS rather than the pREP. Would exposure to MTS training facilitate equivalence responding on a subsequent pREP test?

Experiment 3

Procedure

All training and testing procedures and criteria were as outlined in the General Procedure section. Subjects were trained to criterion using the MTS training procedure. They were then presented with the pREP test, and if necessary retrained and retested until they passed or demonstrated a stable, incorrect performance. They were then presented with MTS tests until they passed or demonstrated stability, and finally were exposed to the pREP test again.

Results and Discussion

Table 1 shows the individual subject data for Experiment 3 (see Appendix C for a detailed breakdown). Subjects 1 and 5 required only 80 training trials to reach the training criterion, and Subjects 2, 3, and 4 required only 120. Of the 5 subjects, 4 produced both symmetry and equivalence responding on the pREP test, on the first (S2), second (S1 and S4), or third (S3) exposures. Subject 5 produced symmetry responding, but not equivalence, even after seven exposures to the test. On the MTS test all 5 subjects produced both symmetry and

equivalence, after just one (S2 to S5) or two (S1) exposures to the test. On the final pREP test, Subject 5 again failed to produce equivalence responding, but responded in accordance with the symmetry relations. All other subjects passed the entire test.

The results of Experiment 3 indicate that when subjects are exposed to MTS training, rather than the pREP training, they are more likely to produce equivalence responding on a subsequent pREP test. Nevertheless, one subject did still fail to produce equivalence on the pREP test, even though equivalence responding subsequently emerged on a MTS test. Furthermore, this subject again failed to produce equivalence on the final exposure to the pREP test. The data from this 1 subject again support the conclusion that MTS more readily produces equivalence, relative to the pREP. Perhaps, therefore, equivalence responding on the pREP might be made even more likely if subjects were *first* exposed to both MTS training *and* testing before the pREP test. Experiment 4 examined this issue.

Experiment 4

Procedure

The training and testing procedures and stability criteria were identical to those used in the previous three experiments. Subjects were trained to criterion using the MTS training procedure. They were then exposed to a MTS test, and retrained and retested until they passed or demonstrated a stable, incorrect performance. They were then presented with pREP tests until they passed or demonstrated stability, and finally were reexposed to the MTS test.

Results and Discussion

Table 1 shows the individual subject data for this experiment (see Appendix D for a detailed breakdown). Subject 5 required only 40 training trials to reach criterion, and the other subjects required up to 440. Subjects 1, 2, 3, and 5 produced both symmetry and equivalence responding on the MTS test on their first exposures to it, but Subject 4 produced neither symmetry nor equivalence responding even after four exposures. On the pREP test, all subjects produced symmetry, but Subjects 3 and 4 failed to produced equivalence responding. When reexposed to the MTS test, Subject 4 again failed to produce both symmetry and equivalence, but all the other subjects passed the entire test.

The results of this experiment are similar to those of Experiment 3, in that one of the subjects produced equivalence on a MTS test, but failed when exposed to the pREP. In contrast to Experiment 3, in which 4 subjects passed the pREP, only 3 subjects did so in the current experiment. This appears to suggest that MTS training *and* testing actually suppresses equivalence responding on the pREP. However, because 1 of the subjects in the current experiment failed both types of test, we are left with only 4 subjects with which we can compare

performances across tests. In doing so, we find that 4 subjects passed using the MTS test, and 3 passed using the pREP test. The results of Experiments 3 and 4 appear broadly similar, therefore.

General Discussion

The results of the four experiments reported here demonstrate that although the pREP can produce equivalence responding, it is not very effective in doing so without some use of MTS procedures as well. Over all four experiments, 10 of the 20 subjects demonstrated equivalence responding on pREP tests, and 16 of the 20 subjects demonstrated equivalence responding on MTS tests. However, the majority of subjects who demonstrated equivalence responding on pREP tests had previously been exposed to MTS training and/or testing procedures.

In Experiment 1, when subjects were exposed to pREP testing immediately after pREP training, only 1 subject produced equivalence responding on the pREP tests. In Experiment 2, pREP training was followed by MTS tests, which were then followed by pREP tests. In this case 2 subjects demonstrated equivalence responding on the pREP tests. Experiments 3 and 4 both used MTS training procedures. In Experiment 3, the pREP tests were presented immediately after MTS training, and produced equivalence responding in 4 subjects. In Experiment 4, pREP tests were presented after exposure to both MTS training and testing. Three subjects demonstrated equivalence responding on the pREP tests but, as was noted earlier, 1 of the subjects who failed to demonstrate equivalence responding on the pREP test had also failed to demonstrate it on the MTS test. It would seem therefore that some exposure to MTS procedures facilitates the production of equivalence responding using pREP procedures.

When exposed to pREP training, subjects, across the four experiments, required an average of 332 training trials to reach criterion, but when exposed to MTS training they needed an average of 156 training trials to reach criterion. This discrepancy is perhaps artificial in that four pREP training tasks are needed to establish one conditional discrimination, whereas only two MTS tasks are needed for this purpose. For example, the tasks: $A1 \rightarrow B1/\text{not } B2$, and $A2 \rightarrow B2/\text{not } B1$ in MTS are represented by the four pREP tasks: $A1 \rightarrow B1 \rightarrow \text{press}$, $A1 \rightarrow B2 \rightarrow \text{no press}$, $A2 \rightarrow B2 \rightarrow \text{press}$, and $A2 \rightarrow B1 \rightarrow \text{no press}$. In view of this, it is readily apparent why the mean number of training trials required to reach criterion in the pREP was almost exactly double the number required in the MTS procedure. That is, perhaps each MTS trial trained the selection of one comparison and the simultaneous rejection of a second comparison, thereby cutting down the number of trials required to teach the discrimination relative to the pREP. Future studies might systematically examine rates of learning using MTS and pREP to determine exactly how these two procedures differentially affect the acquisition of conditional discriminations.

A possible criticism of the current series of experiments is that the MTS procedure involved simultaneous presentations of stimuli whereas the pREP involved successive presentations of stimuli. However, preliminary work in the Cork laboratory has shown that substituting a delayed for a simultaneous MTS procedure does not decrease the relative effectiveness of MTS over the pREP. In the context of the current study, therefore, we may conclude that the use of a simultaneous, rather than a delayed, MTS procedure did not significantly affect the overall pattern of results.

Across all four experiments, many subjects who produced equivalence responding required more than one exposure to the training/testing sequences before demonstrating equivalence. This finding is consistent with other studies using either MTS procedures (e.g., Cullinan, Barnes, Hampson, & Lyddy, 1994; Dymond & Barnes, 1994; Saunders, Saunders, Williams, & Spradlin, 1993), or alternative procedures (e.g., Leader et al., 1996). In the current study the experimental tasks were presented in quasi-random order within blocks of trials, with symmetry, and equivalence tasks intermixed. However, in a study by Fields, Adams, Newman, and Verhave (1992), the need for repeated training and testing was significantly reduced by using a simple-to-complex protocol. This involved subjects being trained and tested for symmetry relations, before being trained and tested for transitivity, and equivalence relations.

In the context of the present series of experiments, use of a simple-to-complex protocol might have reduced the need for repeated training and testing of subjects who demonstrated equivalence responding, and it may also have resulted in more subjects overall demonstrating this performance. Considering the fact that subjects in the current experiments seemed to have more difficulty with the pREP training and testing than the MTS, the effects of a simple-to-complex protocol might be even more pronounced in the pREP than the MTS procedure.

Alternatively, there may be some features of the pREP task itself which inhibit the demonstration of equivalence responding. For example, Dube and McIlvane (1996) pointed out that long response latencies are often observed on equivalence trials, particularly in the early stages of testing (e.g., Saunders, Wachter, & Spradlin, 1988), and therefore go/no-go response procedures can result in no-go responses being recorded on trials where response latencies exceed a certain value. In the pREP tests of the experiments reported here, subjects were allowed 5 s in which to respond, by pressing the space bar (go), or not pressing the space bar (no-go). If a subject did not press the space bar during this response interval, the response was recorded as 'no press' and the next trial was presented. When the data for subjects who failed to demonstrate equivalence responding in the current series of experiments is examined, it can be seen that by far the most common pattern of responding was 'no press' for all the equivalence tasks. This finding could be consistent with Dube and McIlvane's analysis and would

suggest that performance on pREP tests might be improved, if the task was modified to allow subjects to terminate each trial by, for example, pressing a particular key on the keyboard (i.e., press or no press, followed by a termination response).

However, the response latency problem can not provide a complete explanation of the very high level of 'no press' responses on pREP equivalence tests. If insufficient response time were the problem then one would expect performance to improve over repeated presentations of the tests. One would also expect to see at least some pressing responses being recorded on each test block. This was not the case. Close inspection of the data reveals that in final exposures to pREP tests, all but two instances of failure were recorded as no press responses on all equivalence and nonequivalence tasks. This is also consistent with some preliminary data from further studies conducted in the Cork laboratory which examined various modifications to the pREP task. Specifically, it was found that allowing subjects to control the duration or the response interval did not improve performance on pREP tests.

Dube and McIlvane (1996) also pointed out another weakness in go/no-go type procedures that may also be relevant in this case. Specifically, during the training of baseline relations responses to *all* sample-comparison combinations are reinforced (e.g., A1 → B1 → PRESS → REINFORCE; A1 → B2 → NO PRESS → REINFORCE). In effect a subject is rewarded (for pressing) when presented with A1 and B1, *and* is rewarded (for not pressing) in the presence of A1 and B2, and thus the derived relations are rendered ambiguous. Consider the following argument. In the current series of experiments the pREP was used to train the following sample-comparison-correct responses: A1 → B1 → PRESS, A2 → B2 → PRESS, A1 → B2 → NO PRESS, A2 → B1 → NO PRESS, A1 → C1 → PRESS, A2 → C2 → PRESS, A1 → C2 → NO PRESS, A2 → C1 → NO PRESS. In subsequent equivalence tests, either PRESS *or* NO PRESS responses can be predicted on the basis of the baseline training. For example, if the test task presented was C1 → A1 (equivalence), the training history of A1 → B1 → PRESS and B1 → C1 → PRESS would predict a response of PRESS (via equivalence), but the training history of A1 → B2 → NO PRESS, and B2 → C1 → NO PRESS would predict a response of NO PRESS (via equivalence). Furthermore, this problem only relates to transitivity/equivalence relations and therefore does not affect the symmetry tasks. This could account for the fact that most subjects in the current study responded in accordance with symmetry relations, but not in accordance with equivalence relations during the pREP tests.

What remains to be explained, however, is why some subjects failed to respond in accordance with equivalence relations on pREP tests, both before *and* after demonstrating equivalence responding on MTS tests. For example, in Experiment 1, Subjects 1, 3, and 5, having been trained using the pREP, each failed to demonstrate equivalence responding on the pREP test which followed. After retraining they were exposed to MTS

tests, on which all three produced equivalence responding. If Dube and McIlvane's analysis were applied in this case then the ambiguous relations derived by the pREP training, which might account for the failure on the pREP tests, should also have resulted in ambiguous performances on the MTS tests. Similarly in Experiment 2, Subject 5 was exposed to pREP training and demonstrated equivalence responding on a subsequent MTS test, but after retraining failed to demonstrate equivalence responding on the pREP test. Also, in Experiments 3 and 4, where subjects received MTS training, 2 subjects (S5, Exp. 3; S3, Exp. 4) demonstrated equivalence responding on MTS tests, and still failed to demonstrate equivalence responding on pREP tests. In all these cases the argument for ambiguous derived relations would seem to break down.

Another possible explanation for the poor performance of subjects on the pREP test might be that the procedure allows for the formation of one large stimulus class. In effect, during the pREP a subject can simply press (or not press) on all trials, but in the context of MTS the subject must choose one comparison and reject the other on each trial. In other words, the MTS task itself serves to partition the stimuli into two sets, but the pREP task does not. Nevertheless, if this interpretation were correct, then one would expect subjects to have the same difficulties with the symmetry tasks on the pREP tests as they had with the equivalence tasks. This was not the case. This leaves us with one more alternative interpretation of the subjects' failure to press during the equivalence test trials, which is that all the equivalence tasks were responded to as "incorrect," simply because the A and C stimuli had never been presented together during the training. In effect, perhaps the pREP readily facilitates symmetry, but not equivalence responding. Discovering why this should be the case will have to await further empirical enquiry.

In fact, one recently reported study employed procedures not dissimilar to the pREP (Reeve, Fields, Varelas, Rosen, Belanich, & Hobbie, 1996). Specifically, Reeve et al. exposed 18 subjects to a stimulus pairing procedure whereby each experimental trial involved one sample and either a positive or negative comparison being presented successively on a computer screen. The subjects were required to press a key marked with the word YES on sample-positive comparison trials, and to press a key marked with the word NO on sample-negative comparison trials. The instructions given to subjects were quite explicit, in that they were told to "discover whether the words go together." They were also given preexperimental "keyboard familiarization" training using the stimulus pairing task with either semantically related words or semantically unrelated words. Using these procedures, along with a simple to complex protocol, and subject control of the duration of stimulus presentation, Reeve et al. found that 10 of the 18 subjects demonstrated responding in accordance with equivalence relations, without any prior exposure to MTS procedures.

The Reeve et al. experimental tasks seem quite similar to those used in the current experiments. However, certain elements of their

procedure would seem to be more likely to facilitate equivalence responding than the procedures used in the current study. Specifically, the use of the words "go together" in the instructions, the labeling of the response keys with the words yes and no, and the preexperimental training using semantic categories, could, separately or in combination, be expected to function as contextual cues for equivalence responding. For example, when a subject pressed in our experiment, it may have functioned as a linear chain such as A1-B1-press. But in the Reeve et al. study, the instruction "goes with" combined with YES/NO, may have functioned as "Yes, A1 does go with B1."

Although the Reeve et al. procedure was clearly more effective in producing equivalence responding than the pREP procedure employed in the current study, the features discussed above very likely account for this disparity. For this reason it would be important for future research to examine the separate and combined effects of these and other features. By systematically examining the relative effects of the two procedures we may attain a better understanding of those variables that control equivalence responding.

References

- CULLINAN, V. A., BARNES, D., HAMPSON, P. J., & LYDDY, F. (1994). A transfer of explicitly and nonexplicitly trained sequence responses through equivalence relations: An experimental demonstration and connectionist model. *The Psychological Record*, 44, 559-585.
- D'AMATO, M. R., & COLOMBO, M. (1985). Auditory matching-to-sample in monkeys (*Cebus apella*). *Animal Learning and Behavior*, 13, 375-382.
- DUBE, W. V., & MCILVANE, W. J. (1996). Some implications of a stimulus control topography analysis for emergent behavior and stimulus classes. In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 197-218). Amsterdam, Holland: Elsevier.
- DYMOND, S., & BARNES, D. (1994). A transfer of self-discrimination response functions through equivalence relations. *Journal of the Experimental Analysis of Behavior*, 62, 251-267.
- FIELDS, L., ADAMS, B. J., NEWMAN, S., & VERHAVE, T. (1992). Interactions among emergent relations during equivalence class formation. *The Quarterly Journal of Experimental Psychology*, 45B, 125-138.
- HAYES, S. C., & BARNES, D. (1997). Analyzing derived stimulus relations requires more than the concept of stimulus class. *Journal of the Experimental Analysis of Behavior*, 68, 235-244.
- LEADER, G., BARNES, D., & SMEETS, P. M. (1996). Establishing equivalence relations using a respondent-type training procedure. *The Psychological Record*, 46, 685-706.
- MARKHAM, M. R., & DOUGHER, M. J. (1993). Compound stimuli in emergent stimulus relations: Extending the scope of stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, 60, 529-542.

pREP test	10	10	10	9	10	9	0	0	0	0	0	0
MTS retrain to crit. 40												
MTS test	10	10	10	10	10	10	0	0	0	0	0	0
MTS retrain to crit. 40												
pREP test	10	10	10	10	10	10	0	0	0	0	0	0
Test tasks	B1-A1	B2-A2	C1-B1	C2-B2	C1-A1	C2-A2	B1-A2	B2-A1	C1-B2	C2-B1	C1-A2	C2-A1
Required response	P	P	P	P	P	P	NP	NP	NP	NP	NP	NP
Subject 5	No. trials											
MTS train to crit. 80												
pREP test	2	6	2	6	7	8	7	0	0	6	1	1
MTS retrain to crit. 40												
pREP test	6	6	4	6	1	1	1	1	2	2	1	0
MTS retrain to crit. 40												
pREP test	8	5	4	7	0	0	0	0	0	0	0	0
MTS retrain to crit. 40												
pREP test	1	3	0	3	0	0	0	0	0	0	0	0
MTS retrain to crit. 40												
pREP test	2	4	2	2	0	0	0	0	0	0	0	1
MTS retrain to crit. 40												
pREP test	10	9	8	10	0	0	0	0	0	0	0	0
MTS retrain to crit. 40												
pREP test	10	10	10	10	0	0	0	0	0	0	0	0
MTS retrain to crit. 40												
MTS test	10	10	10	10	10	9	0	0	0	0	0	1
MTS retrain to crit. 40												
pREP test	10	10	10	10	0	0	0	0	0	0	0	0

*P = Press, NP = No Press.

Appendix D

Detailed Results for Each Exposure to pREP Tests and MTS Tests in Experiment 4

Test tasks	Experiment 4											
	Symmetry				Equivalence		Non-Symmetry			Non-Equiv.		
Required response	B1-A1	B2-A2	C1-B1	C2-B2	C1-A1	C2-A2	B1-A2	B2-A1	C1-B2	C2-B1	C1-A2	C2-A1
Subject 1	*P	P	P	P	P	P	*NP	NP	NP	NP	NP	NP
Subject 1	No. trials											
MTS train to crit. 440												
MTS test	10	10	10	10	10	9	0	0	0	0	0	1
MTS retrain to crit. 40												
pREP test	10	10	9	9	10	10	0	0	0	0	0	0
MTS retrain to crit. 40												
MTS test	10	10	10	10	10	10	0	0	0	0	0	0
Test tasks	B1-A1	B2-A2	C1-B1	C2-B2	C1-A1	C2-A2	B1-A2	B2-A1	C1-B2	C2-B1	C1-A2	C2-A1
Required response	P	P	P	P	P	P	NP	NP	NP	NP	NP	NP
Subject 2	No. trials											
MTS train to crit. 160												
MTS test	10	9	10	10	9	10	0	1	0	0	1	0
MTS retrain to crit. 40												
pREP test	10	9	9	9	10	9	0	1	0	0	0	0
MTS retrain to crit. 40												
MTS test	10	10	10	10	10	10	0	0	0	0	0	0

Test tasks	B1-A1	B2-A2	C1-B1	C2-B2	C1-A1	C2-A2	B1-A2	B2-A1	C1-B2	C2-B1	C1-A2	C2-A1
Required response	P	P	P	P	P	P	NP	NP	NP	NP	NP	NP
Subject 3	No. trials											
MTS train to crit.	160											
MTS test	10	10	10	10	10	10	0	0	0	0	0	0
MTS retrain to crit.	40											
pREP test	8	10	9	10	0	2	0	0	0	0	0	0
MTS retrain to crit.	40											
pREP test	10	10	10	10	0	0	0	0	0	0	0	0
MTS retrain to crit.	40											
MTS test	10	9	9	10	10	10	0	1	1	0	0	0
Test tasks	B1-A1	B2-A2	C1-B1	C2-B2	C1-A1	C2-A2	B1-A2	B2-A1	C1-B2	C2-B1	C1-A2	C2-A1
Required response	P	P	P	P	P	P	NP	NP	NP	NP	NP	NP
Subject 4	No. trials											
MTS train to crit.	280											
MTS test	10	1	8	9	8	1	0	9	2	1	2	9
MTS retrain to crit.	120											
MTS test	0	10	9	10	0	1	10	0	1	0	10	9
MTS retrain to crit.	40											
MTS test	0	10	9	10	0	0	10	0	1	0	10	10
MTS retrain to crit.	40											
MTS test	0	10	10	10	0	0	10	0	0	0	10	10
MTS retrain to crit.	40											
pREP test	3	7	10	8	0	0	1	0	0	0	0	9
MTS retrain to crit.	40											
pREP test	10	10	10	10	0	1	0	2	0	0	0	0
MTS retrain to crit.	40											
pREP test	8	10	10	10	0	0	0	0	0	0	0	0
MTS retrain to crit.	40											
MTS test	3	10	9	10	0	1	7	0	1	0	10	9
Test tasks	B1-A1	B2-A2	C1-B1	C2-B2	C1-A1	C2-A2	B1-A2	B2-A1	C1-B2	C2-B1	C1-A2	C2-A1
Required response	P	P	P	P	P	P	NP	NP	NP	NP	NP	NP
Subject 5	No. trials											
MTS train to crit.	40											
MTS test	10	10	10	10	9	9	0	0	0	0	1	1
MTS retrain to crit.	40											
pREP test	10	10	10	10	0	0	0	0	0	0	0	0
MTS retrain to crit.	40											
pREP test	9	8	10	10	10	8	0	0	0	2	0	0
MTS retrain to crit.	40											
MTS test	10	10	10	10	10	10	0	0	0	0	0	0

*P = Press, NP = No Press.