



## Assessing stimulus equivalence with a precursor to the relational evaluation procedure

Paul M. Smeets<sup>a,\*</sup>, Marije van Wijngaarden<sup>a</sup>, Dermot Barnes-Holmes<sup>b</sup>,  
Veronica Cullinan<sup>c</sup>

<sup>a</sup> Department of Psychology, Leiden University, P.O. Box 9555, 2300 RB Leiden, The Netherlands

<sup>b</sup> National University, Maynooth, Ireland

<sup>c</sup> National University, Cork, Ireland

Received 17 July 2003; received in revised form 17 July 2003; accepted 21 October 2003

### Abstract

Previous research has demonstrated that, after being trained on multiple match-to-sample (MTS) tasks (A–B, B–C), most human adults respond in accordance with symmetry (B–A, C–B) and equivalence (C–A) when measured with MTS tests and with a precursor to the Relational Evaluation Procedure (pREP). The latter procedure involves conditional go/no-go discrimination tasks, requiring subjects to press a bar during a 5 s interval after the successive presentation of two same-class stimuli, and not to press after the presentation of two different-class stimuli (e.g.  $C_i \rightarrow A_i \rightarrow$  press,  $C_i \rightarrow A_j \rightarrow$  no press). The present study was an effort to replicate these findings. The study consisted of five experiments. Very few subjects evidenced pREP symmetry and equivalence unless they had (a) already demonstrated symmetry and equivalence in a MTS test before, or (b) received pREP pretraining with unrelated stimulus pairs and symmetry was tested before equivalence. Failures to show symmetry were always associated with pressing at or close to 50% of these trials. Failures to show equivalence were associated with pressing or not pressing on (almost) all trials. Current findings are similar to those obtained in equivalence studies involving MTS probes permitting the subjects not to respond to the designated comparisons.

© 2003 Elsevier B.V. All rights reserved.

**Keywords:** Stimulus equivalence; Go/no-go successive discrimination; Matching to sample; Test protocol; Adult humans

### 1. Introduction

Studies on stimulus equivalence almost exclusively use symbolic or arbitrary match-to-sample (MTS) tasks. In its simplest form, such a MTS task consists of three physically disparate stimuli, a sample (e.g. A1 or A2) and two comparisons (e.g. B1 and B2). The subject is required to choose one of the comparisons, depending the sample that is used. During training,

the subject receives positive feedback (e.g. good) for choosing comparison B1 when given sample A1, and for choosing comparison B2 when given sample A2 (A1–B1, A2–B2). All other selections (A1–B2, A2–B1) are followed by negative feedback (wrong). Numerous studies on stimulus equivalence have shown that following the training of these and similar B–C relations (B1–C1, B2–C2), verbal humans readily match, without further training, all same-class stimuli: B1–A1, B2–A2, C1–B1, C2–B2 (symmetry), and C1–A1, C2–A2 (equivalence) (Sidman, 1994, 2000). Similar findings have been obtained with other tasks incorporating elements of conditional

\* Corresponding author. Tel.: +31-71-5273679;

fax: +31-71-5273619.

E-mail address: [Smeets@fsw.leidenuniv.nl](mailto:Smeets@fsw.leidenuniv.nl) (P.M. Smeets).

discrimination tasks (for a detailed review, see Fields et al., 1997). These findings led Saunders and Green (1992, p. 239) to conclude that, "... class formation and perhaps equivalence class formation may be the product of *any* procedure that serves to partition a set of stimuli into subsets of stimuli that are substitutable for one another in certain contexts" (italics added). A similar conclusion was reached by Sidman (1994, p. 384), "An equivalence relation is made up of pairs of stimulus events, with no restriction on the nature of the events that make up the pairs."

Research by Cullinan et al. (1998, 2000, 2001) and Fields et al. (1997), however, suggests that these conclusions may not hold when a precursor to the Relational Evaluation Procedure (pREP) is used. In Cullinan's studies, a pREP trial consisted of the presentation of two successively presented stimuli with a go/no-go response requirement (see D'Amato and Colombo, 1985). During pREP training, the subjects received positive feedback for pressing the space bar of the computer when given a pair of designated same-class stimuli (e.g.  $A1 \rightarrow B1, B1 \rightarrow C1$ ) and for not pressing that bar when given a pair of different-class stimuli (e.g.  $A1 \rightarrow B2, B2 \rightarrow C1$ ). All other stimulus–response relations (i.e. pressing after a pair of different-class stimuli or not pressing after a pair of same-class stimuli) were followed by negative feedback. The pREP was also used to test for the emergence of symmetry (e.g.  $B1 \rightarrow A1 \rightarrow$  press,  $B1 \rightarrow C2 \rightarrow$  no press) and equivalence (e.g.  $C1 \rightarrow A1 \rightarrow$  press,  $C2 \rightarrow A1 \rightarrow$  no press). The Cullinan et al. (1998) study consisted of four conditions, two in which the baseline tasks were trained with the pREP (conditions 1 and 2), and two in which these relations were trained with MTS (conditions 3 and 4). Symmetry and equivalence were tested with the pREP and with MTS, the pREP before the MTS in conditions 1 and 3, and after MTS in conditions 2 and 4 (Condition 1: pREP training  $\rightarrow$  pREP test  $\rightarrow$  MTS test; Condition 2: pREP training  $\rightarrow$  MTS test  $\rightarrow$  pREP test; Condition 3: MTS training  $\rightarrow$  pREP test  $\rightarrow$  MTS test; Condition 4: MTS training  $\rightarrow$  MTS test  $\rightarrow$  pREP test). The results of this and the two following studies consistently showed that most subjects evidenced pREP symmetry but not pREP equivalence unless (a) the baseline relations were trained with MTS (Cullinan et al., 1998, 2000), or (b) two response options, "same" and "different," were added,

and the subjects had learned to select "same" (or an equivalent thereof) when given a pair of same-class stimuli, and to select "different" (or an equivalent thereof) when given a pair of different-class stimuli (e.g.  $A1 \rightarrow B1 \rightarrow$  same,  $A1 \rightarrow B2 \rightarrow$  different) (Cullinan et al., 2001). These findings led the authors to conclude that MTS contains features with pre-experimentally established discriminative properties for responding in accordance with the relations of same and difference (Barnes, 1994; Barnes-Holmes and Barnes-Holmes, 2000; Hayes, 1991) and that the pREP, lacking these features, leads to the formation of (nonseparable) compounds (Cullinan et al., 2000; Wulfert et al., 1991).

Although this account fits the data, it is difficult to reconcile with the results of another study by Smeets et al. (2000a). These authors reported that, after receiving training on simple AB and AC discriminations ( $A1B1+/A1B2-$ ,  $A2B2+/A2B1-$ ,  $A1C1+/A1C2-$ ,  $A2C2+/A2C1-$ ), all adult subjects readily demonstrated simple discrimination performances consistent with symmetry ( $B1A1+/B1A2-$ ,  $B2A2+/B2A1-$ ,  $C1A1+/C1A2-$ ,  $C2A2+/C2A1-$ ) and with equivalence ( $B1C1+/B1C2-$ ,  $B2C2+/B2C1-$ ). These findings are quite remarkable given that the stimulus–response relations were very similar to those in the pREP (e.g.  $A1B1 \rightarrow$  pointing,  $A1B2 \rightarrow$  no pointing).

The present study was part of a project to explore whether the weak equivalence generating effect of the pREP could be related to unexplored features of the training and testing procedures. The pREP could be more sensitive than MTS to the type of test protocols that is used. In all of Cullinan's studies, each test block involved symmetry trials mixed among equivalence trials (complex protocol). In the study by Smeets et al. (2000a), symmetry was tested before equivalence (simple-to-complex protocol). Previous research has shown that MTS produces equivalence more readily with a simple-to-complex than with a complex protocol (Adams et al., 1993; Fields et al., 2000; MacDonald et al., 1986; Smeets et al., 1997, 2000b). Perhaps, this applies even more so to the pREP. If not, perhaps the pREP is more sensitive to the training protocol than MTS. In all studies by Cullinan, a linear protocol was used (training A–B, B–C) whereas the Smeets et al. (2000a) study involved a one-to-many protocol (A–B, A–C). Studies by Arntzen and Holth (1997, 2000) have shown that linear type MTS training

produces equivalence less readily than one-to-many type MTS training.

Because these considerations were highly speculative in nature, the decision was made to first replicate conditions 3 and 4 of the Cullinan et al. (1998) study, in which the baseline tasks were trained with MTS. This replication is described in Experiment 1. Would most subjects show pREP symmetry and equivalence? Quite unexpectedly, they did not unless, like in Condition 4 of Cullinan et al. (1998), they had demonstrated these relations already with MTS. Experiment 2 demonstrated that these contrasting findings were not related to the fact that, in Experiment 1, some of the procedures were different from those used by Cullinan et al. (1998). The following three experiments assessed whether the pREP test performances could be improved by familiarization with the pREP task requirements (Experiment 3), the instruction to respond (i.e. press) on half of the trials (Experiment 4), and using a simple-to-complex testing protocol (Experiment 5).

## 2. Experiment 1

This experiment was a modified replication of conditions 3 and 4 of the Cullinan et al. (1998) study.

### 2.1. Subjects

Eight adults, two males and six females, participated. Their ages ranged from 18 to 27 years. The subjects were recruited through personal contacts, had no previous experience with stimulus equivalence research, and were randomly assigned to one of two conditions, four subjects per condition.

### 2.2. Apparatus and setting

All subjects were trained and tested in a room free of distractions. The stimuli were the nonsense syllables ZID, JOM, TYV, CUG, BEH, and XAD and will be represented here by the alphanumeric A1, B1, C1, A2, B2, and C2, respectively. The stimuli were presented on a computer screen, and subjects responded by pressing various marked keys on the key board. Training and testing trials were presented in blocks. After each trial block, the subjects received a short break outside the room while the experi-

menter checked the results. Sessions lasted from 45 to 180 min. Most subjects completed the experiment in one session.

### 2.3. Procedure

#### 2.3.1. General experimental sequence

The program started with the training of two A–B and two B–C MTS tasks. Then the subjects received pREP and MTS tests to assess the emergence of symmetry (B–A, C–B) and equivalence (C–A). The pREP test was presented before the MTS test in Condition 1, and after the MTS test in Condition 2. Each test presentation was followed by a retraining of the baseline tasks.

#### 2.3.2. MTS training

After being seated in front of the computer monitor, the experimenter read the following instruction aloud: “In this part of the experiment, three nonsense syllables will appear on the screen, one at the top, and two at the bottom. Look at the syllable at the top and choose one of the syllables at the bottom. If you want to select the one on the left, press the Z-key. If you want to select the one on the right, press the M-key. Each time you press, you will see a message on the screen saying either ‘good’ or ‘wrong.’ At the end, a message will appear asking you to call the experimenter. I will be waiting outside. Do you have any questions?” Any questions were answered by the experimenter repeating the entire instruction, at which point she left the room.

The A–B relations were trained first. Blocks of 20 trials were used, 10 A1–B1 trials randomly mixed with 10 A2–B2 trials. Each trial started with the presentation of a sample (e.g. A1) at the top of the screen, followed 1 s later by two comparisons (B1 and B2) presented to the left and right of the sample near the bottom of the screen. The locations of the comparisons (left, right) varied unsystematically over trials. All stimuli remained on the screen until the subject pressed the Z- or M-key. Correct responses were followed by a tone and “good” on the screen (1 s), incorrect responses by another tone and “wrong” on the screen (1 s). Training continued until a subject responded correctly on 18/20 trials with no more than one error on each task (e.g. A1–B1). After also the B–C relations were trained (same procedure), all four

tasks were trained together (mixed training). This training involved blocks of 40 trials, 10 on each task and continued until a subject responded correctly on at least 9 trials on each task.

### 2.3.3. MTS test

Immediately before the test, the experimenter read the following instruction: “In this part of the experiment, three nonsense syllables will appear on the screen, one at the top, and two at the bottom. Look at the syllable at the top and choose one of the syllables at the bottom. If you want to select the one on the left, press the Z-key. If you want to select the one on the right, press the M-key. This time, however, you will not get right/wrong messages, so do whatever you think is right. At the end, a message will appear asking you to call the experimenter. I will be waiting outside. Do you have any questions?”

The test consisted of 60 trials, 40 symmetry trials (B1–A1, B2–A2, C1–B1, C2–B2) randomly mixed

with 20 equivalence trials (C1–A1, C2–A2), 10 trials on each task. The procedures were the same as during training but without feedback (1 s blank screen). Criterion was reached if a subject responded correctly on at least eight or more trials on each task. If necessary, this test was presented two more times, each time followed by MTS retraining (see the following).

### 2.3.4. pREP test

This test started with the following instruction, “In this part of the experiment, two nonsense syllables will appear at the center of the screen, one after another. Then there will be a five second pause. During that pause, you have to press the space bar, or not press that bar. You will not get any right/wrong messages, so do whatever you think is right. Do you have any questions?”

The test consisted of 120 trials, 80 symmetry trials (B1 → A1, B1 → A2, B2 → A1, B2 → A2,

Table 1

Number of training trials, number of test exposures, and the results for symmetry and equivalence during the final test exposure in experiments 1–4

Experiment	Condition	Subject	Trials pREP pretraining	Trials MTS training	Number of tests	Results symmetry	Results equivalence	Number of tests	Results symmetry	Results equivalent	
1	1	1	–	140	pREP		MTS				
		2	–	80	3	Fail	Fail	1	Pass	Pass	
		3	–	80	3	Fail	Fail	3	Pass	Fail	
		4	–	100	3	Fail	Fail	3	Pass	Pass	
	2					MTS		pREP			
		5	–	120	3	Pass	Pass	1	Pass	Pass	
		6	–	140	3	Fail	Pass	3	Fail	Pass	
		7	–	100	1	Pass	Pass	1	Pass	Pass	
2		8	–	80	2	Pass	Pass	3	Pass	Fail	
						pREP		MTS			
		9	–	160	3	Pass	Fail	1	Pass	Pass	
		10	–	80	9	Fail	Fail	3	Pass	Pass	
3		11	–	320	4	Fail	Fail	3	Pass	Fail	
		12	–	120	6	Fail	Pass	2	Pass	Pass	
						pREP		MTS			
		13	60	100	2	Pass	Pass	1	Pass	Pass	
3		14	60	120	3	Pass	Fail	1	Pass	Pass	
		15	40	140	3	Pass	Fail	1	Pass	Pass	
		16	60	100	3	Pass	Fail	3	Pass	Fail	
						pREP		MTS			
4		17	80	140	3	Fail	Fail	3	Pass	Fail	
		18	40	80	3	Pass	Fail	1	Pass	Pass	
		19	100	80	3	Pass	Fail	1	Pass	Pass	
		20	40	120	2	Pass	Pass	1	Pass	Pass	

C1 → B1, C1 → B2, C2 → B1, C2 → B2) mixed with 40 equivalence trials (C1 → A1, C1 → A2, C2 → A1, C2 → A2), 10 trials on each task. On each trial, a sample stimulus (e.g. B1) appeared at the center of screen (1 s), the screen then cleared (1 s) and a same-class stimulus (A1) or different-class stimulus (A2) was presented (1 s) followed by a 5 s pause (blank screen). A response was registered correct if, during the pause, a subject pressed the space bar after the presentation of two same-class stimuli (e.g. B1 → A1) or did not press the space bar after the presentation of two different-class stimuli (e.g. B1 → A2). All other responses were registered incorrect. Criterion was reached if a subject responded correctly on eight or more trials of each task. If necessary, this test was presented two more times, each time followed by MTS retraining.

### 2.3.5. MTS retraining

The procedures were the same as during the MTS training except that blocks of 16 trials were used, 8 A–B trials mixed with 8 B–C trials. This training continued until a subject responded correctly on at least 15 trials.

## 2.4. Results and discussion

Table 1 shows the individual subject data obtained in experiments 1–4. Relevant for Experiment 1, the table shows the total number of trials to complete the MTS training, the number of test presentations, and the results of the final presentation of each test in terms of “Pass” or “Fail”. The results of the MTS retraining are not shown because, with very few exceptions, the subjects always responded correctly.

All subjects learned the MTS baseline tasks in 80–140 trials. None of the four subjects demonstrated pREP symmetry or equivalence when, in Condition 1, this test was presented first. During the subsequent MTS test, all four subjects evidenced symmetry, three of whom also demonstrated equivalence. Condition 2 showed that this discrepancy was not related to the number of test exposures per se. When the MTS test was presented first, three subjects responded according to symmetry, and all four according to equivalence. These performances did not improve or even deteriorated during the subsequent pREP test.

In conclusion, the pREP test results were far worse than those reported in the Cullinan et al. (1998) study, in which MTS training produced class-consistent performances (symmetry and equivalence) in most subjects irrespective of the test order.

## 3. Experiment 2

This experiment was conducted to determine if the negative pREP findings in Condition 1 of Experiment 1 could be related to the fact that the procedures were not completely the same as those used by Cullinan et al. (1998). One modification involved the replacement of two of Cullinan et al.’s stimuli (ROG and DAX) by other stimuli (TYV and XAD). These changes were deemed necessary because, in the Dutch language, ROG and DAX are not meaningless stimuli. The second modification involved the initial training of the A–B and B–C tasks. In Cullinan et al.’s study, these tasks were trained together (blocks of 40 trials). In Experiment 1 of the current study, the A–B and B–C relations were trained separately (blocks of 20 trials) before they were trained in a mixed fashion (blocks of 40 trials). This stepwise approach was intended to facilitate the MTS training. Cullinan et al.’s subjects required a mean of 160 trials (40–440) to complete the baseline training whereas our subjects in Experiment 1 required a mean of 105 trials (80–140). The third modification involved the number of trials per block in the MTS retraining, 16 in Experiment 1, and 40 in Cullinan et al.’s study. Finally, in the current study, the subjects received a maximum of three presentations of each test. In Cullinan et al.’s study, each test was repeated until mastery performance was evident (same criterion as in Experiment 1) or until a stable rate of incorrect responses was demonstrated (i.e. the difference between scores on each individual test task, across two blocks of trials, was no more than 2 out of 10). Although most of the modifications were simply intended to keep the length of the sessions within limits and to prevent exhaustion and boredom, the implications thereof could have been more serious than had been anticipated. Therefore, in Experiment 2, the procedures were identical to those in Condition 3 of Cullinan et al. (1998) in which the pREP test was presented before the MTS test. Would most subjects now show pREP symmetry and equivalence?

### 3.1. Method

Four new subjects, two males and two females between 20 and 29 years, participated. The procedures were the same as in Experiment 1, except that (a) stimuli TYV and XAD were replaced by ROG and DAX, (b) the A–B and B–C matching tasks were always trained together (20 A–B trials mixed with 20 B–C trials), (c) the MTS retraining was extended to 40 trials (same as in the initial training), and (d) testing continued until mastery performance or a stable incorrect performance was evident. Sessions lasted from 90 to 180 min. Two subjects completed the experiment in one session. The other two subjects required two sessions.

### 3.2. Results and discussion

Except for the increased number of MTS training trials ( $M = 170$ ), the test results were similar to those in Condition 1 of Experiment 1 (see Table 1). During the pREP test, Subject 9 showed symmetry but not equivalence. Subject 12 responded according to equivalence but showed no symmetry. Subjects 10 and 11 showed no symmetry or equivalence. During the following MTS test, all four subjects evidenced symmetry, three of them also equivalence. These findings indicated that the negative findings in Condition 1 of Experiment 1 were not related to the use of different and perhaps inadequate procedures. In the following experiments, therefore, the procedures were based again on those used in Experiment 1 (Condition 1).

## 4. Experiment 3

This experiment sought to determine if the absence of class-consistent pREP test performances in experiments 1 (Condition 1) and 2 could be related to the fact that, following the completion of the MTS training, the subjects were ill prepared for the much different pREP tasks. Although the subjects seldom asked any questions, the pREP test instruction may have provided inadequate guidance as to how to respond. Would pREP pretraining lead to improved pREP test performances?

### 4.1. Method

Four new subjects, three females and one male, participated. Their ages ranged from 20 to 24 years. The procedures were the same as in Condition 1 of Experiment 1, except that the MTS training was preceded by pREP pretraining.

#### 4.1.1. pREP pretraining

This training started with the following instruction, “In this part of the experiment, two nonsense syllables will appear at the center of the screen, one after another. Then there will be a 5 s pause. During the 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’ or ‘wrong.’ At the end, a message will appear asking you to call the experimenter. I will be waiting outside. Do you have any questions?”

Four new stimuli were used, SIF (X1), KOQ (X2), MYZ (Y1), and VUP (X2). The procedures were the same as during the pREP test (see Experiment 1) except that the subjects received positive feedback (good + tone 1) when pressing after two same-class stimuli or when not pressing after two different-class stimuli, and negative feedback (wrong + tone 2) when pressing after two different-class stimuli or when not pressing after two same-class stimuli. Blocks of 20 trials were used, five on each task (X1 → Y1 → press, X1 → Y2 → no press, X2 → Y2 → press, X2 → Y1 → no press). The training continued until a subject responded correctly on at least 18 trials of a block with no more than one error on a same task. At that point the subjects proceeded to the MTS training (A–B, B–C).

### 4.2. Results and discussion

All four subjects completed the pREP pretraining (X → Y) and MTS training (A–B, B–C) successfully. The pretraining required 40–60 trials, the training 100–140 trials (see Table 1).

All four subjects evidenced pREP symmetry, one also pREP equivalence. The results of the subsequent MTS tests were the same except that two more subjects evidenced equivalence, both of them immediately. In conclusion, the pREP pretraining successfully promoted symmetry, but not pREP equivalence.

## 5. Experiment 4

This experiment investigated if the superiority of the MTS tests could be related to the fact that these tasks required the subjects to respond to 50% of the displayed stimulus pairs. For example, when given a C–A matching trial, subjects had to choose A1 or A2. In contrast, the pREP test permitted the subjects to press after all or none of the stimulus pairs. Inspection of the data in experiments 1–3 indicated that the rates of press and no-press responses differed markedly between symmetry and equivalence trials. All seven subjects who failed to show pREP symmetry (1–4, 10–12), pressed at or close to 50% of these trials. Some of these subjects pressed at chance level during all tasks (e.g. Subject 3), some systematically pressed when given pairs of different-class stimuli and did not press when given pairs of same-class stimuli (e.g. Subject 1), while others systematically pressed after some stimulus pairs and not after other stimulus pairs, irrespective of whether the paired stimuli were of the same class or not (e.g. Subject 2). By contrast, except for Subject 12 who responded according to equivalence, all these subjects pressed (1, 2, 10, 11) or did not press during most if not all equivalence trials (3, 4). Likewise, of the five subjects who showed pREP symmetry but not pREP equivalence, two pressed (8, 16) and three did not press (9, 14, 15) during all or most C → A trials. Would instructing subjects to press during 50% of *all* trials facilitate higher levels of pREP class-consistent responding?

### 5.1. Method

Four new subjects, two females and two males, participated. Their ages ranged from 22 to 31 years. The procedures were the same as in Experiment 3, except that the sentence, “To do well, you should press half the time” was added to the pREP instructions (pre-training and testing).

### 5.2. Results and discussion

The results (see Table 1) were very similar to those in Experiment 3. Three subjects evidenced pREP symmetry, one of whom also demonstrated pREP equivalence. All subjects who did not show

pREP equivalence (almost) always pressed (17, 19) or never pressed during these trials (18). During the subsequent MTS test, all four subjects demonstrated symmetry, three of whom also showed equivalence.

Following the completion of the experiment, subjects 17–19 were given some extra trials, exposed to the data files, and asked why they had not complied with the instruction to press during 50% of the pREP equivalence trials. All three subjects indicated that they had forgotten the instruction (17, 18) or had ignored the instruction because of its interference with their performance (19). Yet, they were surprised to see that they had pressed or not pressed during most equivalence trials. In any case, the instructions had no effect on the pREP test performance relative to Experiment 3.

## 6. Experiment 5

This experiment sought to determine if the pREP test performances could be improved by separating the symmetry trials from the equivalence trials and testing symmetry before equivalence. This sequential arrangement was expected to have two advantages. First, it might help the subjects assess whether their overall rate of press responses approximated the target set by the instruction (50%). Even if a subject had forgotten the instruction and was inclined to systematically press or not to press during the equivalence trials, the fact that these trials were no longer interspersed among symmetry trials would make it obvious that consistent nondifferential responding would be incorrect. Second, studies involving MTS tasks (see Section 1) have shown that equivalence responding can be facilitated by testing symmetry first. Perhaps this also applies to equivalence measured by the pREP.

### 6.1. Method

Two males and five females, between 18 and 23 years, participated. The procedures were the same as in Experiment 4 except that symmetry (pREP: 80 trials; MTS: 40 trials) was tested before the equivalence (pREP: 40 trials; MTS: 20 trials). Each test could be presented three times, each time followed by MTS re-training.

Table 2  
Number of training trials, number of test exposures, and the results for symmetry and equivalence during the final test in Experiment 5

Subjects	pREP pretraining trials	MTS training trials	pREP symmetry		pREP equivalence		MTS symmetry		MTS equivalence	
			Number of tests	Results	Number of tests	Number of results	Number of tests	Results	Number of tests	Results
21	140	140	3	Fail	3	Fail	3	Fail	3	Fail
22	40	80	3	Pass	1	Pass	1	Pass	1	Pass
23	100	80	3	Pass	3	Pass	1	Pass	1	Pass
24	40	140	2	Pass	1	Pass	1	Pass	1	Pass
25	40	100	1	Pass	1	Pass	1	Pass	1	Pass
26	60	160	1	Pass	2	Pass	1	Pass	1	Pass

## 6.2. Results and discussion

One subject failed to learn the MTS baseline tasks. The results of the remaining six subjects are reported in Table 2. These subjects required 40–140 trials to complete the pREP pretraining ( $X \rightarrow Y$ ) and 80–160 trials to complete the MTS training ( $A-B$ ,  $B-C$ ). All subjects pressed on 50% of the pREP symmetry and equivalence trials. One subject (21) failed all tests. She pressed when given different-class stimulus pairs (e.g.  $B1 \rightarrow A2$ ,  $C2 \rightarrow B1$ ,  $C1 \rightarrow A2$ ) and did not press when given same-class stimulus pairs (e.g.  $B1 \rightarrow A1$ ,  $C1 \rightarrow B1$ ,  $C1 \rightarrow A1$ ) during all pREP tests, and systematically selected different-class comparisons during the MTS tests (e.g.  $B1-A2$ ,  $C2-B1$ ,  $C2-A1$ ). The other five subjects passed all pREP and MTS tests. Thus, the “simple-to-complex” test protocol made the pREP just as effective a test for measuring class-consistent performances (symmetry and equivalence) as MTS.

## 7. General discussion

The results of the present study showed that MTS conditional discrimination training did not produce pREP symmetry and equivalence, unless the subjects had received pREP pretraining and a simple-to-complex testing protocol was used. When, in experiments 1 and 2, the conditions were the same as in the Cullinan et al. (1998) study (i.e. no pREP pretraining, complex test protocol), very few subjects evidenced pREP symmetry or equivalence unless they had shown these relations already during a MTS test (Fisher test,  $P = 0.064$ ); yet, even then (Experiment 1, Condition 2), only half the subjects showed pREP

symmetry and equivalence. Introduction of the pREP pretraining in experiments 3 and 4 led to a marked increase of pREP symmetry (Fisher test,  $P = 0.005$ ) but not of pREP equivalence unless, in Experiment 5, also a simple-to-complex test protocol was used (Fisher test,  $P = 0.049$ ).

Furthermore, failures to demonstrate pREP symmetry and equivalence were associated with different response rates. All 10 subjects who failed to demonstrate symmetry pressed at or close to 50% of all symmetry trials. By contrast, with the exception of 1 subject (21) in Experiment 5, all 14 subjects in experiments 1 through 4 who failed to demonstrate equivalence, consistently pressed (8 subjects) or did not press during these trials (6 subjects).

These findings are much different from those obtained by Cullinan et al. (1998) who reported that (a) almost all subjects demonstrated pREP symmetry and equivalence irrespective of whether the subjects had already shown MTS symmetry and equivalence before, and (b) failures to show pREP equivalence were almost always associated with not pressing (eight/nine subjects). For an overview of the major findings obtained by Cullinan et al. (1998) and in the present study, see Table 3.

How can these findings be accounted for? Not by the procedures. In Experiment 2, the procedures were the same as those used by Cullinan, yet the results were very similar to those obtained in Experiment 1. At the present time, it remains unclear why we failed to replicate findings previously produced by our research group. The only obvious difference that remains between Cullinan et al. (1998) study and the current study is that the former was conducted in Ireland and the latter in Holland. However, this seems like a highly unlikely source of discrepancy between the



Table 3  
 Overview of pREP test results in the Cullinan et al. (1998) study and in the current study

	Cullinan et al.'s (1998) study		Current study					
	Complex test protocol		Experiment 1		Experiment 2	Experiment 3*	Experiment 4*	Experiment 5*
	Condition 3*	Condition 4	Complex test protocol		Replication Cullinan et al.'s (1998) Condition 3	Complex test protocol + pREP pretraining	Complex test protocol + pREP pretraining + instruction	S-T-C test protocol + pREP pretraining + instruction
			Condition 1*	Condition 2				
<b>Symmetry</b>								
Number of subjects passing test	5/5	5/5	0/4	3/4	1/4	4/4	3/4	5/6
Number of subjects who consistently pressed	0	0	0	0	0	0	0	0
Number of subjects who consistently did not press	0	0	0	0	0	0	0	0
<b>Equivalence</b>								
Number of subjects passing test	4/5	4/5	0/4	3/4	1/4	1/4	1/4	5/6
Number of subjects who consistently pressed	0	0	2	1	2	1	2	0
Number of subjects who consistently did not press	1	1	2		1	2	1	0

Note. S-T-C: simple-to-complex. Asterisk (\*) indicates the pREP test presented before MTS.

two data sets. Because we have successfully replicated numerous studies across the Dutch and Irish laboratories over many years, we have no reason to suspect that some unidentified cross-laboratory difference is at work here. Perhaps, these discrepancies are related to sample size. The positive pREP findings obtained in Condition 3 of the Cullinan et al. (1998) study were based on 5 subjects while our negative pREP findings in experiments 1–4 were based on 16 subjects. Thus, the original data should be viewed with caution. Despite the questions raised by the failure to replicate, the findings across all five of the current experiments have highlighted some important issues with respect to the pREP and its relationship to the ubiquitous MTS procedure, and it is to these that we now turn.

One might argue that the negative pREP findings resulted from the fact that these tasks were very different from those used for training (MTS). This would explain why in equivalence studies, including the current one, subjects who are trained on MTS baseline tasks, readily show MTS symmetry and equivalence, and why in the Smeets et al. (2000a) study, training of simple AB and AC discrimination tasks readily produced class-consistent simple BC discriminations. This account, however, does not explain why in other studies, pREP training (Cullinan et al., 1998, conditions 1 and 2) and simple exposure to pairs of sequentially presented same-class stimuli (Leader et al., 1996, 2000; Smeets et al., 1997) readily produced MTS symmetry and equivalence, and why pREP training did *not* produce pREP equivalence (Cullinan et al., 1998, conditions 1 and 2; Cullinan et al., 2000, 2001). Collectively, these findings indicate that the difficulties in obtaining pREP equivalence (and symmetry) are not related to the fact that these tasks were different from those used for training (MTS), but to the inadequacy of the pREP for testing derived relations.

Instead, present findings suggest that pREP test performances are sensitive to modifications of (a) the composition of the stimulus pairs, and (b) the functions (sample and comparison) of each element of a pair. When in Condition 2 of Experiment 1 the pREP test was presented last, the subjects were exposed to the same stimulus pairs (albeit now presented in a sequential fashion) that were used in the preceding MTS symmetry and equivalence trials. Also the functions of the stimuli were the same (e.g. C1 as sample and A1 as comparison). In this condition, most

subjects evidenced pREP symmetry and equivalence (hence pressed on 50% of the trials) even though they had not been exposed to pREP tasks before (no pre-training). When, in Condition 1 of Experiment 1 and in Experiment 2, the pREP tests were presented first, the stimulus pairs used for testing symmetry were the same as those during MTS training except that the functions (sample, comparison) of the pair members were changed. Again, the subjects responded to 50% of the trials but without evidencing symmetry, unless they had received pREP pretraining (experiments 3–5). When the stimulus pairs *and* the functions of the pair members were different from those used in the MTS baseline tasks (pREP equivalence), the subjects responded nondifferentially and systematically pressed or did not press (experiments 1–4), unless the conditions discouraged them from doing so (Experiment 5).

Interestingly, similar negative test findings have been reported in a small number of equivalence studies involving MTS probes with an added default option (“none,” “can’t answer”), or in which subjects are given the opportunity to skip test trials (Duarte et al., 1998; Innis et al., 1998; Smeets et al., 2000c). In the study by Smeets et al. (2000c), for example, subjects were given sheets showing baseline tasks (A–B, B–C) mixed with symmetry (B–A, C–B) and equivalence tasks (C–A). Some of the subjects received these sheets with the instruction to skip “impossible to solve” tasks. Other subjects did not receive this instruction. Most subjects in the no-instruction condition completed all tasks and evidenced symmetry and equivalence. By contrast, most subjects in the instruction condition responded to all baseline tasks but skipped most if not all symmetry and equivalence tasks. Similar findings have been reported by Duarte et al. (1998) who used MTS probes with an added “can’t answer” option. The pREP offers a similar default option by allowing the subjects to press or not to press during all trials. Thus, it would be interesting to see if our current pREP test findings, notably those of experiments 1 through 4, would be much different from those obtained with MTS tests with an added default option (e.g. not to respond within a 5 s interval).

Although the present findings indicate that the pREP is a less adequate procedure for measuring stimulus class formation than MTS (Cullinan et al., 1998, 2000, 2001), they also show that this problem can be easily remedied. Most subjects showed symmetry

after pREP pretraining, and most of them also showed pREP equivalence when the simple-to-complex test protocol was used. These findings are consistent with and extend previous work by Fields and colleagues who reported that the emergence of complex MTS performances can be substantially enhanced by the prior induction and the demonstration of easier and/or prerequisite performances (Adams et al., 1993; Fields et al., 2000). Future research, therefore should assess whether similar findings will be obtained when the baseline tasks are trained with pREP rather than with MTS.

## References

- Adams, B.J., Fields, L., Verhave, T., 1993. Effects of test order on intersubject variability during equivalence class formation. *Psychol. Rec.* 43, 133–152.
- Arntzen, E., Holth, P., 1997. Probability of stimulus equivalence as a function of training design. *Psychol. Rec.* 47, 309–320.
- Arntzen, E., Holth, P., 2000. Equivalence outcome in single subjects as a function of training structure. *Psychol. Rec.* 50, 603–628.
- Barnes, D., 1994. Stimulus equivalence and relational frame theory. *Psychol. Rec.* 44, 91–124.
- Barnes-Holmes, D., Barnes-Holmes, Y., 2000. Explaining complex behavior: two perspectives on the concept of generalized operant classes. *Psychol. Rec.* 50, 251–265.
- Cullinan, V.A., Barnes, D., Smeets, P.M., 1998. A precursor to relational evaluation procedure: analyzing stimulus equivalence. *Psychol. Rec.* 48, 121–145.
- Cullinan, V.A., Barnes-Holmes, D., Smeets, P.M., 2000. A precursor to the relational evaluation procedure: analyzing stimulus equivalence II. *Psychol. Rec.* 50, 467–492.
- Cullinan, V.A., Barnes-Holmes, D., Smeets, P.M., 2001. A precursor to the relational evaluation procedure: searching for contextual cues that control equivalence responding. *J. Exp. Anal. Behav.* 76, 339–349.
- D'Amato, M.R., Colombo, M., 1985. Auditory matching-to-sample in monkeys (*Cebus apella*). *Anim. Learn. Behav.* 13, 375–382.
- Duarte, A.M., Eikeseth, S., Rosalez-Ruiz, J., Baer, D.M., 1998. The effects of a can't-answer response option and instructions on stimulus equivalence. *Psychol. Rec.* 48, 631–646.
- Fields, L., Reeve, K.F., Varelas, A., Rosen, D., Belanich, J., 1997. Equivalence class formation using stimulus-pairing and yes-no responding. *Psychol. Rec.* 47, 661–686.
- Fields, L., Varelas, A., Reeve, K.F., Belanich, J., Wadhwa, P., DeRosse, P., Rosen, D., 2000. Effects of prior conditional discrimination symmetry transitivity, and equivalence testing on the emergence of new equivalence classes. *Psychol. Rec.* 50, 443–466.
- Hayes, S.C., 1991. A relational control theory of stimulus equivalence. In: Hayes L.J., Chase P.N. (Eds.), *Dialogues on Verbal Behavior*. Context Press, Reno, NV, pp. 19–40.
- Innis, A., Lane, S.D., Miller, E.R., Critchfield, T.S., 1998. Stimulus equivalence: effects of a default-response option on emergence of untrained stimulus relations. *J. Exp. Anal. Behav.* 70, 87–102.
- Leader, G., Barnes, D., Smeets, P.M., 1996. Establishing equivalence relations using a respondent-type training procedure. *Psychol. Rec.* 46, 685–706.
- Leader, G., Barnes-Holmes, D., Smeets, P.M., 2000. Establishing equivalence relations using a respondent-type training procedure III. *Psychol. Rec.* 50, 63–78.
- MacDonald, R.P.F., Dixon, L.S., LeBlanc, J.M., 1986. Stimulus class formation following observational learning. *Anim. Learn. Behav. Disabil.* 6, 73–87.
- Saunders, R.R., Green, G., 1992. The nonequivalence of behavioral and mathematical equivalence. *J. Exp. Anal. Behav.* 57, 227–241.
- Sidman, M., 1994. *Equivalence Relations and Behavior: A Research Story*. Authors Cooperative, Boston.
- Sidman, M., 2000. Equivalence relations and the reinforcement contingency. *J. Exp. Anal. Behav.* 74, 127–146.
- Smeets, P.M., Barnes, D., Cullinan, V., 2000a. Establishing equivalence classes with match-to-sample format and simultaneous-discrimination format conditional discrimination tasks. *Psychol. Rec.* 50, 721–744.
- Smeets, P.M., Barnes-Holmes, D., Nagle, M., 2000b. Transfer and stimulus equivalence classes derived from simultaneously presented S+ and S- stimuli. *Eur. J. Behav. Anal.* 1, 33–49.
- Smeets, P.M., Dymond, S., Barnes-Holmes, S., 2000c. Instructions, stimulus equivalence and stimulus sorting: effects of sequential testing arrangements and a default option. *Psychol. Rec.* 50, 339–354.
- Smeets, P.M., Leader, G., Barnes, D., 1997. Establishing stimulus classes in adults and children using a respondent-type training procedure: a follow-up study. *Psychol. Rec.* 47, 285–308.
- Wulfert, E., Dougher, M.J., Greenway, D.E., 1991. Protocol analysis of the correspondence of verbal behavior and equivalence class formation. *J. Exp. Anal. Behav.* 56, 489–504.