The Analysis of Verbal Behavior 1997, **14**, 57-83

# Relating Equivalence Relations to Equivalence Relations: A Relational Framing Model of Complex Human Functioning

**Dermot Barnes and Neil Hegarty**
**National University of Ireland, Cork**

AND

**Paul M. Smeets**
**Leiden University**

The current study aimed to develop a behavior-analytic model of analogical reasoning. In Experiments 1 and 2 subjects (adults and children) were trained and tested for the formation of four, three-member equivalence relations using a delayed matching-to-sample procedure. All subjects (Experiments 1 and 2) were exposed to tests that examined relations between equivalence and non-equivalence relations. For example, on an equivalence-equivalence relation test, the complex sample B1/C1 and the two complex comparisons B3/C3 and B3/C4 were used, and on a nonequivalence-nonequivalence relation test the complex sample B1/C2 was presented with the same two comparisons. All subjects consistently related equivalence relations to equivalence relations and nonequivalence relations to nonequivalence relations (e.g., picked B3/C3 in the presence of B1/C1 and picked B3/C4 in the presence of B1/C2). In Experiment 3, the equivalence responding, the equivalence-equivalence responding, and the nonequivalence-nonequivalence responding was successfully brought under contextual control. Finally, it was shown that the contextual cues could function successfully as comparisons, and the complex samples and comparisons could function successfully as contextual cues and samples, respectively. These data extend the equivalence paradigm and contribute to a behaviour-analytic interpretation of analogical reasoning and complex human functioning, in general.

---

When a number of related conditional discriminations are explicitly trained, the stimuli that enter into these discriminations may become related to each other in ways that were not explicitly trained. Suppose, for example, that choosing the arbitrary stimulus B is reinforced in the presence of the sample stimulus A, and choosing another arbitrary stimulus C is also reinforced in the presence of A. Following this explicit training, a subject may, in the absence of further reinforcement, respond in accordance with; (i) the symmetry relations, given B select A, and given C select A, and (ii) the combined symmetry and transitivity relations, given B select C, and given C select B. Responding in accordance with symmetry and transitivity, in a matching-to-sample context, is normally accepted as evidence that the three stimuli (A, B, C) participate in an equivalence relation (see Barnes & Holmes, 1991; Fields, Adams, Verhave, & Newman, 1990; Sidman, 1992).

In the investigation of stimulus equivalence researchers have generally focused on the relations between simple or single-element stimuli. A number of relatively recent studies, however, have started to examine emergent performances using complex or multi-element stimuli (e.g., Markham & Dougher, 1993; Stromer & Stromer, 1990a, 1990b). In the first study by Stromer and Stromer, they trained rela-

Data from Experiments 1 and 2 were presented during the symposium: "Emerging Themes in Stimulus Equivalence," at the Annual Conference of the Experimental Analysis of Behaviour Group, London, England, U.K. (April 12, 1995). The first author would like to acknowledge the doctoral research of Gina Lipkens as an important source of inspiration for the current set of experiments. Requests for reprints should be sent to Dermot Barnes, Behavior Analysis and Cognitive Science Unit, Department of Applied Psychology, National University of Ireland, Cork, Ireland.

tions of the form AB-D and AC-E, and obtained evidence for emergent relations among all possible pairs of single stimuli (e.g., A-B, D-B, B-C, B-E, and D-E). In the second study, Stromer and Stromer trained the relations A-C, B-D, and AB-E, and again found strong evidence for emergent relations among all possible pairs of stimuli (A-D, B-C, C-E, and D-E). These data clearly showed that it is possible to train human subjects in a series of related conditional discriminations using compound or multi-element sample stimuli, and that subjects often respond in systematic ways during non reinforced probe trials to single elements of the compound samples used during training.

These analyses were extended in a more recent study (Markham & Dougher, 1993), in which subjects were trained in a number of related conditional discriminations using multi-element and single-element samples and single-element comparisons. For the nonreinforced test trials, the researchers reconfigured the individual elements used in training to form new multi-element and single-element stimuli. In Experiment 3, for example, five subjects were trained in nine AB-C relations and three C-D relations and were then probed for the emergence of equivalence (D-AB) and for AD-B and BD-A relations. Three out of five subjects demonstrated the emergence of all tested relations; one subject failed to show equivalence, but demonstrated AD-B and BD-A relations, and the remaining subject showed neither.

Another recent study (Perez-Gonzalez, 1994) that used multi-element stimuli adopted a somewhat different strategy to that employed in previous behavior-analytic studies of emergent performances. This study determined whether relations among sample and comparison stimuli, that had been established in prior conditional discrimination training, would control selection of comparisons in a new task (Perez-Gonzalez, 1994). The basic procedure was as follows. First, relations between particular stimuli were established by training A-B relations through conditional discriminations (A1-B1, A2-B2, and A3-B3). Second, an analogue of the "yes/no" response in the presence of particular relations was achieved with pairs of sample stimuli; the members of each pair of sample stimuli had previously been related as sample-correct comparison or sample-incorrect comparison (e.g., A1-B1: sample-correct comparison or A1-B2: sample-incorrect comparison). The analogues of "yes" and "no" were two novel comparison stimuli, X1 and X2. During training, responses to X1 were reinforced if the two stimuli in the sample had participated in the sample-correct comparison relation (e.g., A1-B1), and responses to X2 were reinforced if the two stimuli in the sample had participated in the sample-incorrect comparison relation (e.g., A1-B2).

The next stage of the study then examined transfer of the relational control described above. Relations between novel stimuli were trained (P1-Q1, P2-Q2, and P3-Q3), those stimuli were then presented as paired samples (e.g., P1-Q1 or P1-Q2), and X1 and X2 were presented as the comparisons in a test (hereafter referred to as the PQX test). Results showed that subjects chose the stimulus X1 when a sample contained the sample-correct comparison relation from the previous training (e.g., P1-Q1), whereas subjects chose X2 when a sample contained the sample-incorrect relation (P1-Q2) from the training. In effect, the PQX test demonstrated that selections of X1 and X2 transferred from the explicit training with the A-B stimuli to the new PQ pairs of stimuli, without explicitly training the X1 and X2 functions ("Yes" and "No," respectively) for the PQ stimulus pairs.

Interestingly, this approach to the study of learning to form and then subsequently respond to arbitrary relations may also help to develop a behaviour-analytic interpretation of advanced reasoning abilities in humans (e.g., Gick & Holyoak, 1983). Consider, for example, the following question based on the classic proportion scheme (A : B: : C : ?); "apple is to orange as dog is to; (i) sheep, or (ii) book?" If apple and orange participate in an equivalence relation in the context "fruit," and "dog" and "sheep" participate in an equivalence relation in the context "animals," then we would expect a person to pick "sheep" as the correct answer. In effect, the response would be in accordance with the derived equivalence relation between two already established separate equivalence relations (see Figure 1, upper section). The three experiments reported here attempted to model this form of "reasoning" behavior (i.e., equivalence-equivalence responding) in the operant laboratory using both adults and children. Demonstrating this complex type of stimulus control would; (i) extend the basic stimulus equivalence effect to include equivalence-equivalence relations, (ii) supplement the recent work on stimulus equivalence using multi-element stimuli, and (iii) support and extend the "Yes/No" analysis of responding to conditional relations developed by Perez-Gonzalez (1994) (see below).

Before outlining the first experiment in the current study we should explain that our use of the term equivalence-equivalence responding (or relating equivalence relations to equivalence relations) is purely descriptive, and is not offerred as a new technical term. We take the view that equivalence-equivalence responding is an example of a relational network as defined by relational frame theory (e.g., Hayes, 1991, 1994; Barnes & Holmes, 1991; Barnes, 1994). From this perspective, equivalence responding, and what we call equivalence-equivalence responding, are viewed as instances of overarching or generalized operant behavior that is produced by multiple-exemplar training provided, in large part, by the verbal community. Relational frame theory has been described in detail in many previous publications, and thus we will not cover old ground here. However, two very recent articles (Barnes & Roche, 1996; Hayes & Barnes, 1997) are perhaps the most relevant to the current study, and the interested reader should consult these to understand fully our rationale for using the current terminology.

Experiment 1 examined the relations between two separate equivalence relations and between two separate nonequivalence relations. Subjects were first trained and tested for the formation of four, three-member equivalence relations (i.e., train; A1→B1, A1→C1, A2→B2, A2→C2, A3→B3, A3→C3, A4→B4, A4→C4, and test; B1↔C1, B2↔C2, B3↔C3, B4↔C4). After successfully passing the equivalence test, subjects were tested to determine whether they would relate pairs of stimuli to other pairs of stimuli based on their participation in equivalence relations. In effect, subjects were presented with samples that contained two stimuli that were from one derived equivalence relation (e.g., B1C1), and were given the opportunity to choose comparisons that contained two stimuli that were from a second, separate derived equivalence relation (e.g., B3C3). If subjects relate equivalence relations to equivalence relations, this will represent an important extension to the study reported by Perez-Gonzalez in which subjects responded "yes" (i.e., chose X1) to an explicitly trained sample-correct comparison relation. In the current study, subjects were also presented with samples that contained two stimuli that participated in *separate*, derived equivalence relations (e.g., B1C2), and were given the opportunity to choose comparisons that contained two stimuli that participated in other, separate, derived equivalence relations (e.g., B3C4). If subjects relate nonequivalence relations to nonequivalence relations this will represent a further extension to the study reported by Perez-Gonzalez in which subjects responded "no" (i.e., chose X2) to an explicitly trained sample-incorrect comparison relation. In the interests of clarity, Experiments 2 and 3 of the current study will be introduced at a later point in the paper.

## GENERAL METHOD

### *Subjects*

One non-psychology undergraduate (male, aged 21), nine non-psychology graduates (four males and five females, aged between 23 and 35), one nongraduate adult (male, aged 24), and two children (both males, aged 9 and 12) participated as sub-

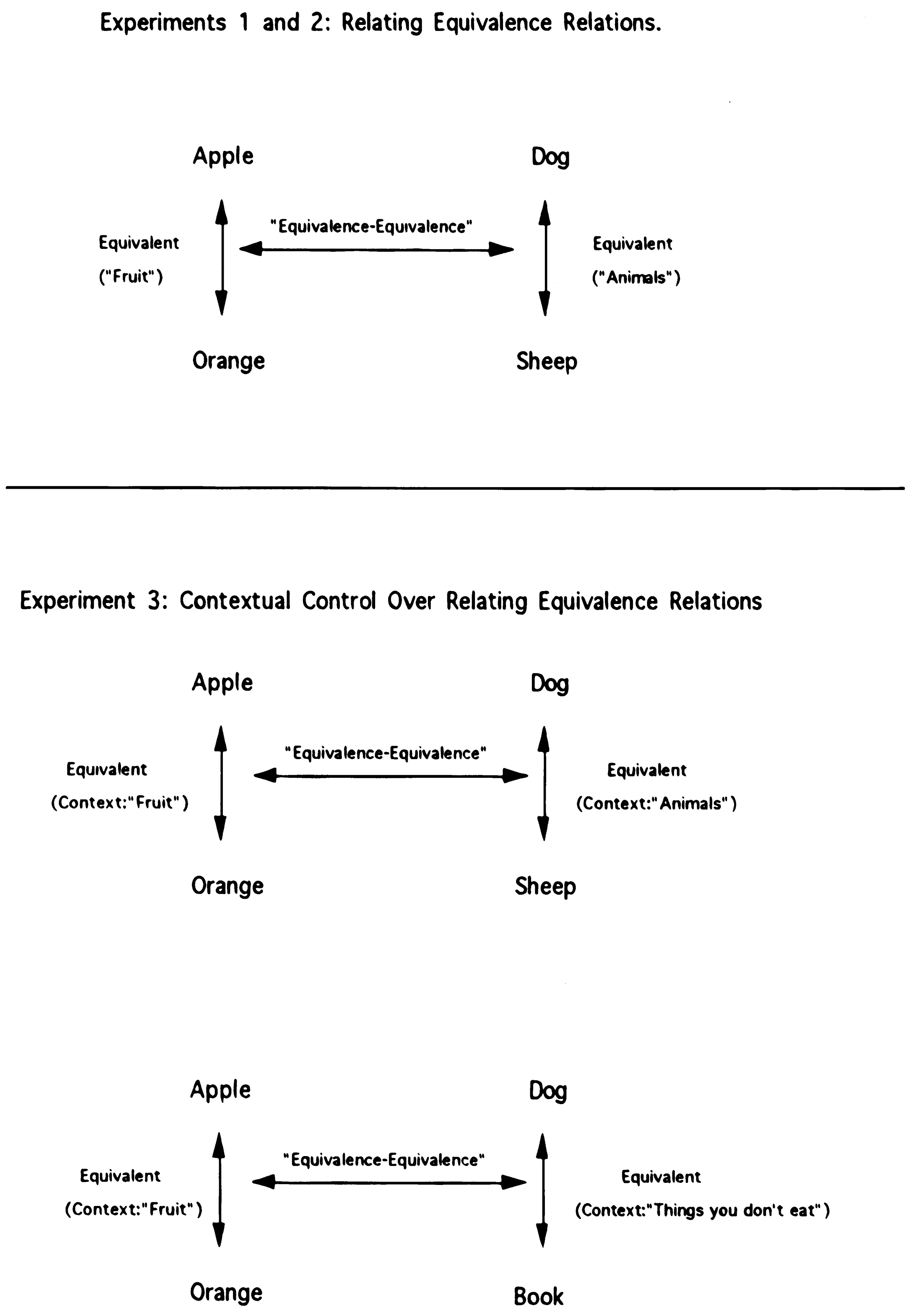


Fig. 1. Schematic representation of the types of equivalence relations that might underlie complex reasoning abilities.

jects across the three experiments. Subject 6 in Experiment 1 and Subject 6 in Experiment 2 were also used in Experiment 3, but they were numbered 2 and 3, respectively, in the third experiment (Subject 1 in Experiment 3 was experimentally naive). All subjects were recruited through personal contacts. No payment was given for participation, but the children were given sweets at the end of each session, irrespective of their performances. Before the experiment began subjects were told that the experiment involved working on a computer based task for at least one, 1 hr session, and were then asked if they still wished to continue. One person refused at this point to participate. All subjects were exposed to the experimental procedures individually. If a subject did not complete the experiment in one session, he or she was asked to return on a subsequent day (usually the following day). To ensure that the previously established performances were still intact, at the beginning of the next session the subject was re-exposed to those stages of the experiment that he or she had previously completed. On some occasions, therefore, a subject could successfully complete a particular stage in the experiment (see procedure sections), but would be reexposed to that stage for a second time. All subjects were asked not to discuss their participation in the experiment with anyone, and sessions were arranged so that subjects did not meet each other in the vicinity of the laboratory. When the experiment was finished each subject was thanked and fully debriefed.

*Apparatus and Stimuli*

An Apple Macintosh® PC with a 14-inch monitor, that displayed black characters on a white background, was used to present stimuli and record data during the experiment. Each subject was seated before the computer in an small experimental room. The stimuli were twelve nonsense syllables (ZID, CUG, VEK, YIM, BEH, DAX, ROG, PAF, MAU, JOM, KIB, FUB). Each nonsense syllable was designated A1, B1, C1, A2, B2, C2, A3, B3, C3, A4, B4, and C4. The alphanumeric designations are for descriptive purposes only and were not shown to the subjects. The nonsense syllables were assigned randomly to their designated roles for each subject (e.g., ZID may have functioned as A1 for one subject, but as A2 for another). In Experiment 3, two additional stimuli were used (these were XXX and OOO) and were designated Ct1 and Ct2 (XXX and OOO were randomly assigned to their roles as Ct1 and Ct2 for each subject).

*Delayed Matching-to-Sample*

A delayed, arbitrary matching-to-sample procedure was used for training and testing. On any given trial, the sample appeared in the center of the screen. When the subject pressed the "G" key (marked with a white paper dot) on the computer keyboard, the sample was removed, and the screen remained blank for .4 seconds; four comparison stimuli were then presented, one in each corner of the screen. The subject selected one of the comparisons by pressing one of four letter keys, R, C, U, N (each marked with a white paper dot), with each key corresponding to one of the four corners of the screen (i.e., R-top left, C-bottom left, U-top right, and N-bottom right; all other keys were disabled). When one of these four keys was pressed, the screen cleared, and during training phases of the experiments, responses defined as correct (see below) produced the word "CORRECT" on the monitor together with a 0.5 s beep. Choosing any of the other three comparisons produced the word "WRONG" and no accompanying sound from the computer. The locations of the correct and incorrect comparisons was counterbalanced across trials. After 2 s, the feedback message (i.e., "Correct" or "Wrong") was removed from the screen and after a further 2 s intertrial interval a new trial began. During testing no feedback was presented, and the computer simply proceeded directly to the intertrial interval.

On some tasks, sample and comparison stimuli were composed of pairs of nonsense syllables; each pair was presented side-by-side separated by a 3mm space (e.g., CUG ZID). On these trials, a pair of nonsense syllables was presented in the center of the screen as a sample, and when the subject emitted the sample-observing response (i.e., pressed the "G" key) the sample was removed, and .4 sec. later a further two pairs of stimuli were presented as two comparisons. One pair was presented in the bottom-left hand corner, and the other pair appeared in the bottom-right hand corner of the screen (the locations of the "correct" and "incorrect" comparisons were counterbalanced across trials). The subject selected the comparison on the left by pressing the "C" key and selected the comparison on the right by pressing the "N" key (i.e., all other keys were disabled). These tasks, on which pairs of nonsense syllables were presented as samples and comparisons, occurred only during the test phases of an experiment (i.e., no feedback was presented for responding on these tasks).

On yet other tasks (Experiment 3) a contextual stimulus was presented 1 cm above the sample stimulus (see Experiment 3 for a detailed explanation of the contextual stimuli). The contextual stimulus remained on the screen throughout each matching-to-sample trial. That is, it remained on screen in the presence of the sample, during the .4 sec. sample-comparison delay, and in the presence of the comparison stimuli. When a subject selected one of the comparison stimuli (again, locations were counterbalanced across trials), the contextual cue and the comparisons were removed from the screen, and the appropriate feedback and/or the intertrial interval followed.

*Procedure: Experiment 1*

At the beginning of each session a subject was presented with the following instructions displayed on the computer screen:

> You must look at the nonsense syllable in the center of the screen, press the marked center key, and then choose one of the four nonsense syllables that appear at each corner of the screen, by pressing one of the four marked keys on the keyboard (sometimes there will be only two choices at the bottom of the screen).
>
> To choose the top-left syllable press the marked key on the top-left.
>
> To choose the top-right syllable press the marked key on the top-right.
>
> To choose the bottom-left syllable press the marked key on the bottom-left.
>
> To choose the bottom-right syllable press the marked key on the bottom-right.
>
> Sometimes the computer will give you feedback and sometimes it will not.
>
> PRESS THE SPACE-BAR TWICE TO CONTINUE

After the subject had read the instructions, the experimenter left the room.

*Matching-To-Sample Training*

During Experiment 1 (see Figure 2), eight delayed matching-to-sample tasks were used to train the subjects in a series of related conditional discriminations (i.e., A1 as sample, pick B1 as comparison, A2-B2, A3-B3, A4-B4, A1-C1, A2-C2, A3-C3, A4-C4). The following training sequence was employed with three of the six subjects (1, 2, and 3). The four A-B tasks were presented in a quasi-random order in blocks of four trials (i.e., with the constraint that each of the four A-B tasks was presented once within each block), until a subject produced a minimum of eight consecutively correct responses. The four A-B and four A-C tasks were then presented in a quasi-random order in blocks of four trials (with the constraint that each of the eight tasks was presented once across every two successive four-trial blocks), until a subject produced a minimum of 12 consecutively correct responses. The subjects then proceeded to the equivalence test. The same training sequence was used for the remaining three subjects (4, 5, and 6), except that the training commenced with the A-C stimulus pairs, and subsequently introduced the A-B pairs.

*Equivalence Testing*

During the equivalence test, eight different matching-to-sample tasks were used. Four of the tasks presented one of the B stimuli as a sample and the four C stimuli as comparisons, and the other four tasks presented one of the C stimuli as a sample and the four B stimuli as comparisons. These eight tasks tested the following

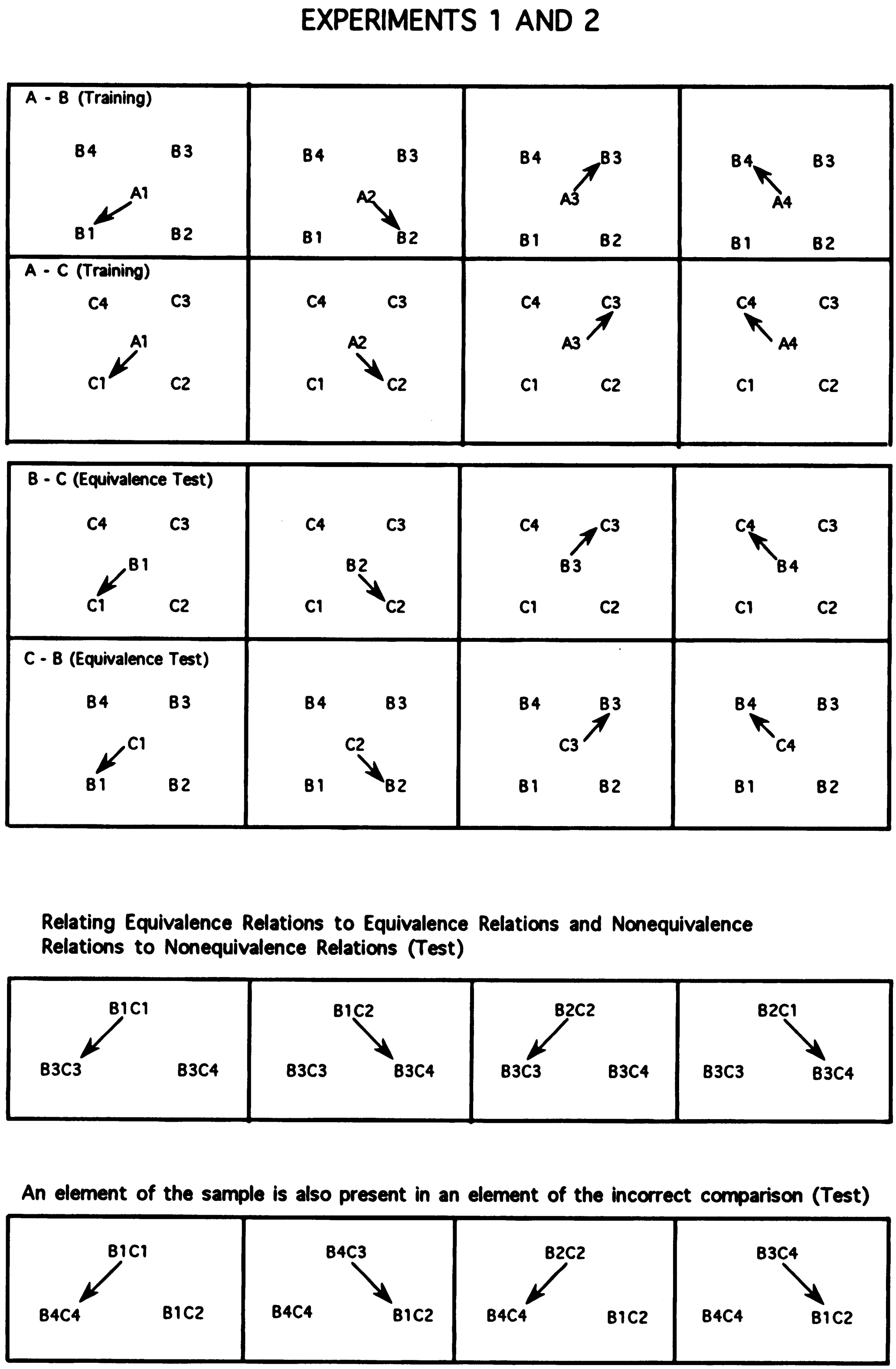


Fig. 2. Schematic representation of the matching-to-sample tasks that were used to train and test for equivalence relations, and equivalence-equivalence relations.

equivalence relations; B1-C1, B2-C2, B3-C3, B4-C4, C1-B1, C2-B2, C3-B3, and C4-B4 (see Figure 2). The eight equivalence tasks were presented, without feedback, in a quasi-random order in a single block of 40 trials (i.e., each of the eight tasks was presented five times within the 40 trial block). If a subject produced at least 4 out of 5 correct responses on each of the eight tasks this performance was defined as equivalence responding. If, however, the subject failed to demonstrate equivalence, he or she was returned to the matching-to-sample training described above. Following retraining, the subject was reexposed to the equivalence test, and if necessary retrained and retested until he or she produced equivalence responding.

*Equivalence-Equivalence Testing*

During the equivalence-equivalence test, subjects were presented with one pair of nonsense syllables as a sample and two pairs of nonsense syllables as comparisons. These pairs of syllables were either from the same equivalence relations (e.g., B1C1) or from two *separate* equivalence relations (e.g., B1C2). For the first block of 20 trials in the equivalence-equivalence test (see Figure 2) subjects were presented, in a quasi-random order, with four different matching-to-sample tasks (i.e., each task presented five times within a 20 trial block). Each task presented a different sample stimulus (i.e., B1C1, B1C2, B2C2, and B2C1), but the same two comparison stimuli were presented on each of the four tasks (i.e., B3C3 and B3C4). It was predicted, that following the presentation of a sample stimulus that contained two equivalent nonsense syllables (e.g., B1 and C1), subjects would choose the comparison that contained a further two equivalent nonsense syllables (e.g., in the presence of B1C1 a subject should choose B3C3 rather than B3C4). In effect, subjects should relate one equivalence relation to a second, separate equivalence relation (i.e., B1 and C1 participate in one equivalence relation, and B3 and C3 participate in another equivalence relation). It was also predicted, that following the presentation of a sample stimulus that contained two nonequivalent nonsense syllables (e.g., B1 and C2), subjects would choose the comparison that also contained a further two nonequivalent nonsense syllables (e.g., in the presence of B1C2 a subject should choose B3C4 rather than B3C3). In effect, subjects should relate one nonequivalence relation to a second, separate nonequivalence relation (i.e., B1 and C2 participate in one nonequivalence relation, and B3 and C4 participate in another nonequivalence relation). (In the interests of brevity, relating equivalence relations to equivalence relations, and relating nonequivalence relations to nonequivalence relations will simply be described as equivalence-equivalence responding, and where appropriate the generic term equivalence-equivalence relation/s will also be used).

For the second block of 20 trials in the equivalence-equivalence test (see Figure 2) subjects were presented, in a quasi-random order, with a further four different matching-to-sample tasks (i.e., each task presented five times within a 20 trial block). Each task presented a different sample stimulus (i.e., B1C1, B4C3, B2C2, B3C4), but the same two comparison stimuli were presented on each of the four tasks (i.e., B4C4 and B1C2). It was predicted, that following the presentation of a sample stimulus that contained two equivalent nonsense syllables (e.g., B1 and C1), subjects would choose the comparison that contained a further two equivalent nonsense syllables (e.g., in the presence of B1C1 a subject should choose B4C4 rather than B1C2). In effect, subjects should relate one equivalence relation to a second, separate equivalence relation. It was also predicted, that following the presentation of a sample stimulus that contained two nonequivalent nonsense syllables (e.g., B4 and C3), subjects would choose the comparison that also contained a further two nonequivalent nonsense syllables (e.g., in the presence of B4C3 a subject should choose B1C2 rather than B4C4). It should be noted that we deliberately designed the second block of testing tasks so that each incorrect comparison contained an element that was also present in the sample stimulus. We assumed that presenting a sample and an incorrect comparison that both contained an element in common would generate competition between responding in accordance with arbitrary equivalence relations and responding in accordance with the non-arbitrary relation of reflexivity, and thus more errors might be observed in the second test block relative to the first.

Because it was uncertain whether the predicted performances would, in fact, be the most likely outcome, a nonpredicted stability criterion was employed for the equivalence-equivalence testing. That is, a subject's performance was defined as stable when he or she selected the *same, but not necessarily correct*, comparison stimulus on each of the eight tasks at least four out of five times across a single exposure to the two 20 trial blocks. In effect, a subject could produce a stable, but nonpredicted (incorrect) performance on the equivalence-equivalence test. If a subject did not produce a stable performance on the equivalence-equivalence test, he/she was reexposed to the entire experimental sequence for a second time (i.e., matching-to-sample training, equivalence testing, and equivalence-equivalence testing). It was decided at the beginning of the study that this recursive training and testing procedure would be continued until; (i) the subject successfully completed the matching-to-sample training and equivalence testing, and produced a stable (but not necessarily correct) performance on the equivalence-equivalence test, or (ii) the subject completed five exposures to the entire experimental sequence, without producing a stable performance on the equivalence-equivalence test. When a subject produced a stable performance on the equivalence-equivalence test, and sufficient time was available, he or she was reexposed to the entire experimental sequence (Subjects 1 and 4 could not complete this final requirement).

The reader should note that a large number of matching-to-sample tasks could have been used at this stage to test for equivalence-equivalence relations. Nevertheless, we used only those tasks that were deemed necessary to demonstrate the predicted derived relations. For example, we did not include any complex stimuli that contained two elements that had entered into a single conditional discrimination (e.g., A1B1, B1C1, etc.). This strategy was adopted because unexpected and/or uncontrolled sources of stimulus control are more likely to occur as the number of testing tasks increases (Steele & Hayes, 1991). Furthermore, using an exhaustive set of tasks may have "overworked" at least some of the subjects during the experiment, and thus increased the likelihood of "failure" caused simply by fatigue and/or boredom.

## Results and Discussion

Figures 3 and 4 present the total number of training trials per exposure, the percentage of correct responses made during each exposure to the equivalence test, and the percentage of correct responses made during each exposure to the equivalence-equivalence test. A detailed breakdown of the subjects' performances can be found in the appendix.

Because Subject 1 was a 12 year old boy, his data will be described first and in the most detail; the reader should note, however, that this subject was actually the last subject to be exposed to the procedures of Experiment 1. This subject required a total of 184 training trials. He then failed the equivalence test (i.e., 22.5% correct). After a further 208 training trials, he improved his performance on the equivalence test, but still failed to demonstrate clear equivalence responding (87.5% correct). Following another 60 training trials, the subject again failed to pass the equivalence test (92.5% correct), but this time he failed by only one response (i.e., on one block of five trials the subject emitted three, rather than four, correct responses out of five; see appendix). During his fourth exposure to the experimental sequence, he passed the equivalence test (after 20 training trials). The subject was then exposed to the equivalence-equivalence test, and he produced a 57.5 percent correct performance that was also defined as unstable (he did not choose the same comparison at least four times on each of the eight tasks). He was again exposed to equivalence training (20 trials), equivalence testing (95% correct), and equivalence-equivalence testing. On this final exposure to the equivalence-equivalence test, the subject achieved 92.5 percent correct and also met the stability criterion (he chose the "correct" comparison at least four times on each of the eight tasks).

The remaining six subjects from Experiment 1 also completed the matching-to-sample training, demonstrated equivalence responding, and produced the predicted equivalence-equivalence responding. Subject 2 required a total of 360 training trials, and was provided with five exposures to the equivalence test, and three exposures to the equivalence-equivalence test. A summary for subjects 3 to 6 is as follows: Subject 3, 180 training trials, 6 equivalence tests, and 2 equivalence-equivalence tests; Subject 4, 264 training trials, 4 equivalence tests, and 2 equivalence-equivalence tests; Subject 5, 84 training trials, 3 equivalence tests, and 3 equivalence-equivalence tests.; Subject 6, 268 training trials, 4 equivalence tests, and 4 equivalence-equivalence tests (due to experimenter error, Subject 6 was exposed to the equivalence-equivalence test before successfully passing the equivalence test).

The reader should note that after Subjects' 2, 3, 5, and 6 had successfully passed the equivalence-equivalence test, they were each exposed once more to the matching-to-sample training, to the equivalence test, and to the equivalence-equivalence test (see final three columns, including "Train_" on their respective graphs, Figures 3 and 4). These additional exposures were conducted to determine whether all of the training and test performances would remain relatively stable. All four subjects successfully retrained and passed the equivalence, and equivalence-equivalence tests during these re-exposures.

The data from Experiment 1 clearly showed that a range of subjects, including a 12 year old boy, could successfully relate

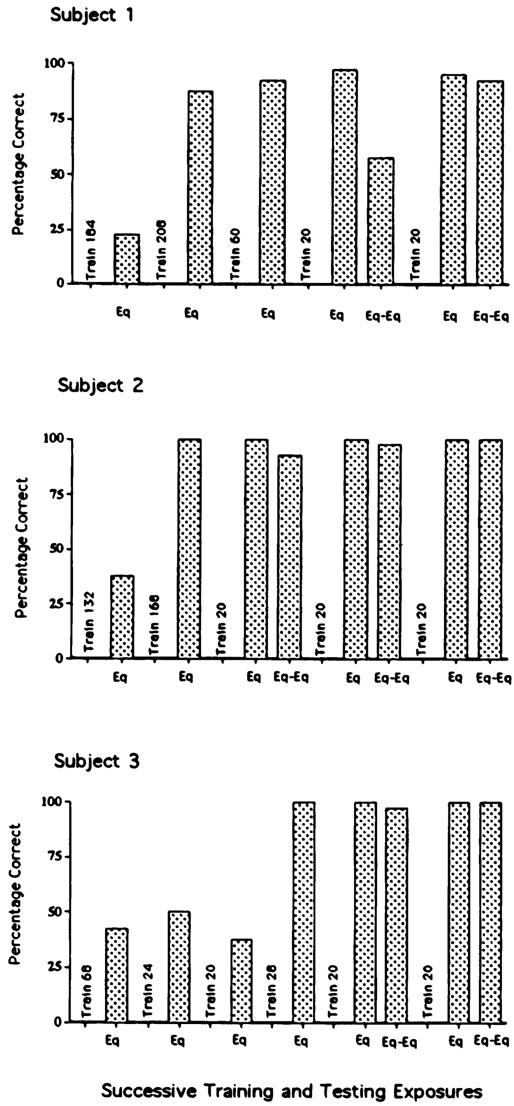


Fig. 3. Percentage of correct responses across successive exposures to the training and testing stages of Experiment 1 for Subjects 1 to 3. Eq indicates a standard equivalence test, and Eq-Eq indicates those tests that examined equivalence-equivalence relations.

equivalence relations to other, separate equivalence relations (e.g., B3C3 was matched to B1C1) and non-equivalence relations to other, separate non-equivalence relations (e.g., B3C4 was matched to B1C2) in the absence of explicit reinforcement. One issue that arises from Experiment 1, however, relates to the fact that subjects were required to pass the standard equivalence test before being exposed to the equivalence-equivalence test. It is unclear, therefore, to what extent passing

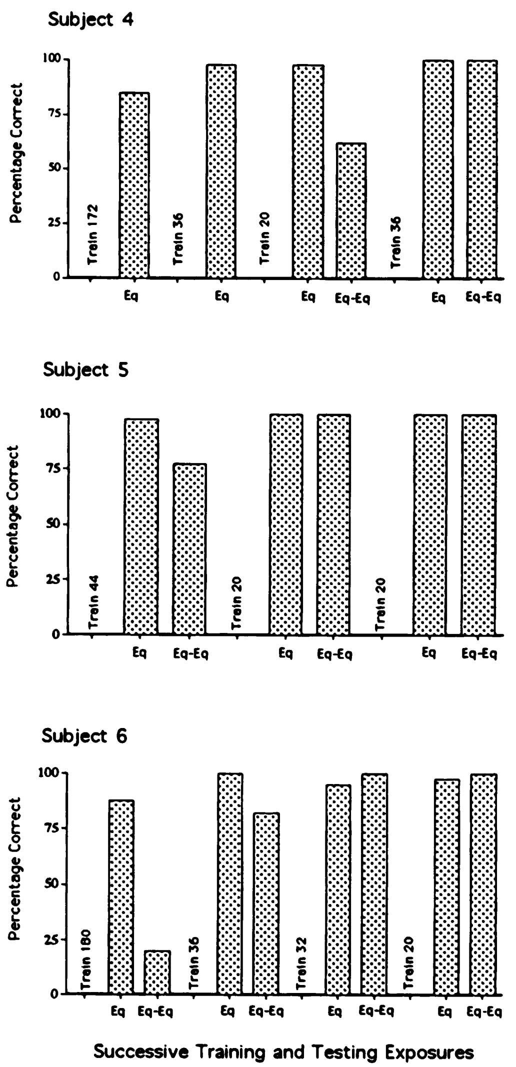


Fig. 4. Percentage of correct responses across successive exposures to the training and testing stages of Experiment 1 for Subjects 4 to 6. Eq indicates a standard equivalence test, and Eq-Eq indicates those tests that examined equivalence-equivalence relations.

the standard equivalence test contributed towards a subject's performance on the equivalence-equivalence test. It may be, for example, that successful exposure to the equivalence test facilitates passing the equivalence-equivalence test, in the same way that testing for equivalence appears to faciliate passing a transfer of functions test (Wulfert & Hayes, 1988; Barnes, Browne, Smeets, & Roche, 1995). Experiment 2 addressed this issue. Six naive subjects were trained and tested using the same procedures employed in Experiment 1, but were exposed to the equivalence-equivalence test (until they produced a stable performance) *before* being exposed to the standard equivalence test.

## EXPERIMENT 2

### METHOD

*Procedure: Experiment 2*

The procedures of Experiment 2 were the same as those employed in Experiment 1, except that subjects were exposed to the equivalence-equivalence test before being exposed immediately, and without further training, to the standard equivalence test. Subjects were required to produce a stable but not necessarily correct performance on the equivalence-equivalence test (i.e., choose the same comparison at least four times out of five), before proceeding to the standard equivalence test.

### RESULTS AND DISCUSSION

Figures 5 and 6 present the total number of training trials per exposure, the percentage of correct responses made during each exposure to the equivalence test, and the percentage of correct responses made during each exposure to the equivalence-equivalence test (see appendix for a detailed breakdown).

Because Subject 1 was a 9 year old boy, his data will be described first and in the most detail; again the reader should note that this subject was actually the last subject to be exposed to the procedures of Experiment 2. On his first exposure to the conditional discrimination training, Subject 1 required a total of 136 trials. During his subsequent exposure to the equivalence-equivalence test he produced an unstable performance (i.e., he did not emit the same four, or five, responses on each of the eight tasks). Furthermore, only 45 percent of his responses on this test were correct. During his second exposure a further 60 training trials were required, and although the performance was still unstable the number of correct responses improved slightly to 55 percent. On his third exposure he required a further 40 training trials, and his performance on the equivalence-equivalence test improved again to 67.5 percent correct, but remained unstable. Only 20 training trials were required during his fourth exposure, but the number of correct responses decreased to 52.5 percent, and the performance was still unstable. On the fifth exposure the subject required 36 training trials, and the number of correct responses improved to 77.5 percent, but once again the performance was unstable. Twenty four training trials were required during the sixth exposure, and the subject then produced 100% correct responding on the equivalence-equivalence test (by definition, the performance was also stable). He was then exposed to the standard equivalence test, without further training, and emitted 100% correct responding.

The remaining five subjects from Experiment 2 also completed the matching-to-sample training, produced stable equivalence-equivalence responding, and immediately demonstrated standard equivalence responding (i.e., they passed the standard equivalence test on the first exposure). Subject 2 required a total of 348 training trials, and was provided with five exposures to the equivalence-equivalence test, and one exposure to the equivalence test. A summary for subjects 3 to 6 is as follows: Subject 3, 428 training trials, 4 equivalence-equivalence tests, and 1 equivalence test; Subject 4, 144 training trials, 3 equivalence-equivalence tests, and 1 equivalence test; Subject 5, 248 training trials, 3 equivalence-equivalence tests, and 1 equivalence test; Subject 6, 140 training trials (approximately), 2 equivalence-equivalence tests, and 1 equivalence test The reader should note, that during Subject 6's first exposure to the training there was a general power failure in the laboratory, and thus an exact record of the number of training trials was lost. Nevertheless, the experimenter estimated the number of trials completed before the power failure to be in the region of 100, and this figure was added to the 20 trials that the subject needed to complete the training during the next exposure (i.e., approximately 120 training trials were required before the subject was exposed to the equivalence-equivalence test).

The data from Experiment 2 showed that a new group of subjects, including a 9 year old boy, successfully related equivalence relations to other, separate equivalence relations and related nonequivalence relations to other, separate nonequivalence relations *before* being exposed to a standard equivalence test. In effect, passing the standard equivalence test was not a necessary prerequisite for passing the equivalence-equivalence test.

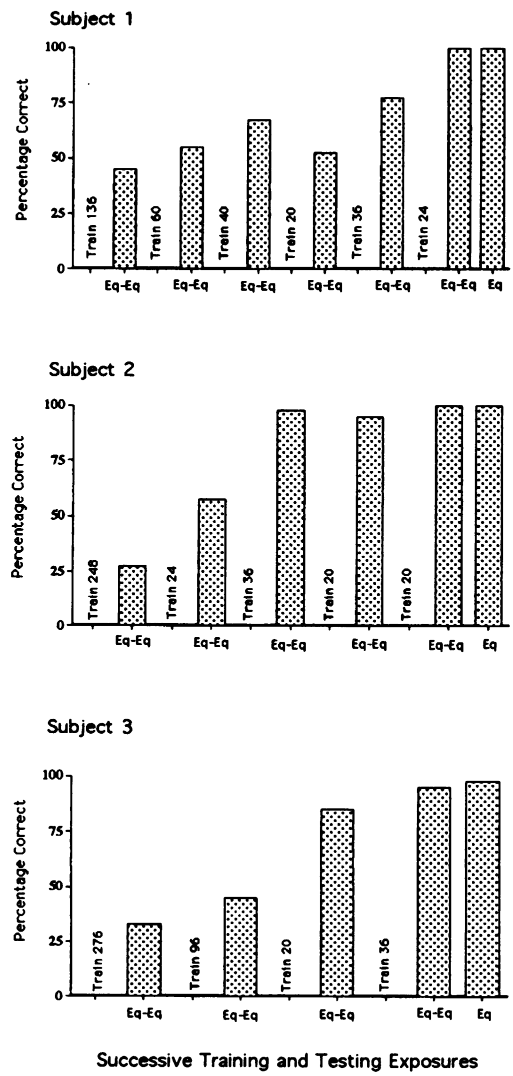


Fig. 5. Percentage of correct responses across successive exposures to the training and testing stages of Experiment 2 for Subjects 1 to 3. Eq indicates a standard equivalence test, and Eq-Eq indicates those tests that examined equivalence-equivalence relations.

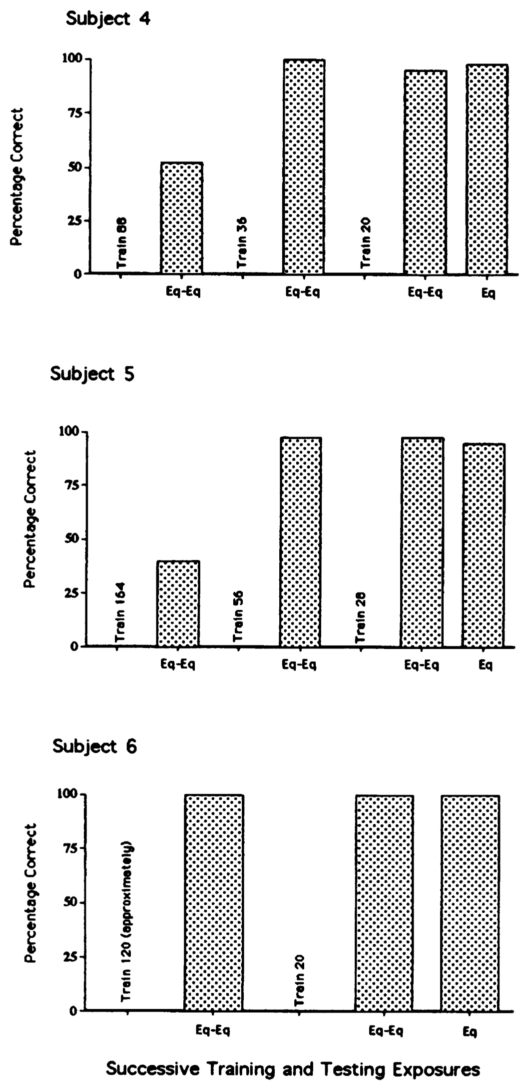


Fig. 6. Percentage of correct responses across successive exposures to the training and testing stages of Experiment 2 for Subjects 4 to 6. Eq indicates a standard equivalence test, and Eq-Eq indicates those tests that examined equivalence-equivalence relations. Due to a general power failure in the laboratory, the first set of training trials completed by Subject 6 had to be estimated (see text for details).

## EXPERIMENT 3

Experiments 1 and 2 demonstrated that both adult and younger subjects may respond in accordance with equivalence-equivalence relations in the absence of explicit reinforcement. One reason for investigating this pattern of behavior was to provide the necessary data to support a behavior-analytic interpretation of human reasoning abilities. Experiment 3 attempted to extend the scope of the current behavior analysis of reasoning to include the role of context. For illustrative purposes, recall the example given earlier (Figure 1, top panel), based on the classic proportion scheme (A : B: : C : ?); "apple is to orange as dog is to; (i) sheep, or (ii) book?" As outlined before, if apple and orange participate in the equivalence relation "fruit," and "dog" and "sheep" participate in the equivalence relation, "animals," then we would expect a person to pick "sheep" as the correct answer. However, "dog" and "sheep" participate in an equivalence relation only in certain contexts (e.g., within the context of "animals"). If the context for the second part of the proportion scheme were to change, for example, to "things you don't eat," then "dog" and "sheep" would be nonequivalent (at least for most Irish people) and "dog" and "book" would participate in the new equivalence relation (i.e., "book" is the correct answer; see Figure 1, bottom panel). This type of contextual control over human reasoning was examined in Experiment 3, using the relational framing model developed in the two previous experiments.

Experiments 1 and 2 both trained and tested for equivalence, and equivalence-equivalence relations without establishing explicit contextual control over these relations. A number of equivalence studies have shown, however, that it is possible to control the stimuli that participate in equivalence relations by presenting a contextual stimulus during training and testing (e.g., Bush, Sidman, & deRose, 1989; Wulfert & Hayes, 1988; Wulfert, Greenway, & Dougher, 1994). For example, if a subject was trained in four conditional discriminations in the presence of two contextual stimuli (i.e., Contextual stimulus 1; A1-B1, A1-C1, A2-B2, A2-C2: Contextual stimulus 2; A1-B1, A1-C2, A2-B2, A2-C1), it is likely that four contextually controlled equivalence relations would emerge (i.e., Contextual stimulus 1, A1-B1-C1, A2-B2-C2 : Contextual stimulus 2, A1-B1-C2; A2-B2-C1). The primary aim of Experiment 3 was to demonstrate this form of contextual control over equivalence-equivalence relations.

Subjects were trained and tested for the formation of eight contextually controlled equivalence relations (i.e., Context 1, A1-B1-C1, A2-B2-C2, A3-B3-C3, A4-B4-*C4*: Context 2; A1-B1-C1, A2-B2-C2, A3-B3-*C4*, A4-B4-*C3*). In effect, the relations remained the same as in Experiments 1 and 2, except that in the presence of Context 2, C3 and C4 "swapped" equivalence relations. If contextual control was established over the equivalence-equivalence relations, subjects should choose, for example, B3C3 when presented with the sample B1C1 in the presence of Context 1, but should choose B3C4 when presented with the same sample in the presence of Context 2.

A subsidiary aim of Experiment 3 was to extend the analysis of contextually controlled equivalence-equivalence relations. In the introduction, we described a study by Markham and Dougher (1993) that examined a number derived relations that emerge when subjects are trained on a matching-to-sample task using compound or multi-element stimuli. The findings from this study raised an interesting possibility concerning the role of contextual control over equivalence responding (Markham & Dougher 1993, p. 540-541). Specifically, these researchers suggested that using compound samples may have caused elements from these samples to function as contextual stimuli. For example, when subjects were trained in the following relations; A1B1-C1, A2B2-C3, and A3B1-C2, it is possible that A1 and A2 might have functioned as contextual stimuli for the conditional functions of B1, or perhaps B1 might have functioned as a contextual stimulus for the conditional functions of the A stimuli. More importantly, for present purposes, however, is the fact that these researchers also showed that both elements from the A and B samples could function independently as comparisons during test trials. In effect, their data suggested that a contextual stimulus might also function successfully as a comparison stimulus (see Wulfert et al., 1994, for a similar suggestion). With this idea in mind, the subsidiary aim of Experiment 3 (in the current study) was to determine whether contextually controlled equivalence-equivalence responding would be maintained if the samples were presented as contextual stimuli, the comparisons were presented as samples, and the contextual stimuli were presented as comparisons. For example, would subjects choose Ct1 when presented with the sample B3C3 in the presence of B1C1, but choose Ct2 when presented with the sample B3C4 in the presence of B1C1? As an aside, although this performance could be described in terms of the contextual stimuli functioning as separable elements of a complex sample, we will continue to use the relational frame concept of contextual stimuli (see Dymond & Barnes, 1995, for evidence to support the use of relational frame terminology over that of separable stimulus compounds).

## METHOD

*Procedure: Experiment 3*

As outlined previously, three subjects participated in Experiment 3; Subject 1 was experimentally naive, whereas Subjects 2 and 3 had successfully completed Experiments 2 and 3 respectively. The procedure for Experiment 3 was similar to Experiments 1 and 2, with the following exceptions (the reader should note that the instructions used in the previous two Experiments were also used in Experiment 3, although they made no reference to the presentation of the contextual stimuli).

*Matching-To-Sample Training*

During Experiment 3, sixteen delayed matching-to-sample tasks were used to train the subjects in a series of related conditional discriminations in the presence of two contextual stimuli; Ct1; A1-B1, A2-B2, A3-B3, A4-B4, A1-C1, A2-C2, *A3-C3*, *A4-C4*, and Ct2; A1-B1, A2-B2, A3-B3, A4-B4, A1-C1, A2-C2, *A3-C4*, *A4-C3* (XXX and OOO were randomly assigned to their roles as Ct1 and Ct2 respectively). In the presence of Ct1 the same conditional discriminations were trained as in Experiments 1 and 2; in the presence of Ct2, however, one of the conditional discriminations was reversed, in that subjects were trained to choose C4 when presented with the A3 sample and to choose C3 when presented with the A4 sample (see underlined numerics above and Figure 7).

The eight A-B tasks (i.e., four in the presence Ct1 and four in the presence of Ct2) were presented in a quasi-random order in blocks of eight trials, with the constraint that each of the eight tasks was presented once within each block, until a subject produced a minimum of eight consecutively correct responses. Subsequently, the eight A-C tasks (i.e., four in the presence Ct1 and four in the presence of Ct2) were presented in a quasi-random order in blocks of eight trials, with the constraint that each of the eight tasks was presented once within each block, until a subject produced a minimum of eight consecutively correct responses. Finally, all sixteen A-B and A-C tasks were presented in a quasi-random order in blocks of 16 trials, with the constraint that each of the sixteen tasks was presented once within each sixteen trial block, until a subject produced a minimum of 16 consecutively correct responses. Subjects then proceeded to the equivalence tests.

*Equivalence Test with Contextual Stimuli*

The equivalence test with contextual stimuli employed sixteen different matching-to-sample tasks (see Figure 7). Eight of the tasks presented the contextual stimulus, Ct1, and the remaining eight presented contextual stimulus, Ct2. In the presence of Ct1, the following eight equivalence relations were tested; B1-C1, B2-C2, B3-C3, B4-C4, C1-B1, C2-B2, C3-B3, and C4-B4. In the presence of Ct2, the remaining eight equivalence relations were tested; B1-C1, B2-C2, B3-C4, B4-C3, C1-B1, C2-B2, C3-B4, and C4-B3 (see Figure 7). During the first 40 trials of this test, the eight B-C tasks (four in the presence of Ct1, and four in the presence of Ct2) were presented, without feedback, in a quasi-random order (i.e., each of the eight tasks was presented five times in a single block of 40 trials). The second block of 40 test trials presented the eight C-B tasks (four in the presence of Ct1, and four in the presence of Ct2), without feedback, in a quasi-random order (i.e., each of the eight tasks was presented five times in a single block of 40 trials). If a subject produced at least 4 out of 5 correct responses on each of the sixteen tasks, this performance was defined as contextually controlled equivalence responding.

*Equivalence-Equivalence Test with Contextual Stimuli*

Subjects were exposed immediately to the equivalence-equivalence test, whether or not they had demonstrated contextually controlled equivalence responding. In the presence of the contextual stimuli, subjects were presented with one pair of nonsense syllables as a sample and two pairs of nonsense syllables as comparisons (see Figure 8). These pairs of syllables were from either the same equivalence relation (e.g., B1C1) or from two separate equivalence relations (e.g., B1C2). During the equivalence-equivalence test subjects were presented, in a quasi-random order, with eight matching-to-sample tasks without feedback (i.e., each task presented five times within a 40 trial block). (Because the introduction of contextual stimuli doubled the number of testing tasks required to examine the predicted derived relations, the tasks from the second block of equivalence-equivalence test trials used in the previous experiments were not used in Experiment 3).

Each task presented either Ct1 or Ct2 with one of four different sample stimuli (i.e., B1C1, B1C2, B2C2, and B2C1), and the same two comparison stimuli (i.e., B3C3 and B3C4). It was predicted that following the presentation of a sample stimulus that, *in the presence of the contextual stimulus*, contained two equivalent nonsense syllables, subjects would choose the comparison that contained a further two equivalent nonsense syllables (e.g., in the presence of Ct1 and given the sample B1C1, subjects should choose B3C3 rather than B3C4, but in the presence of Ct2, and the same sample, subjects should choose B3C4 rather than B3C3). It was also predicted that following the presentation of a sample stimulus that, *in the presence of the contextual stim-*

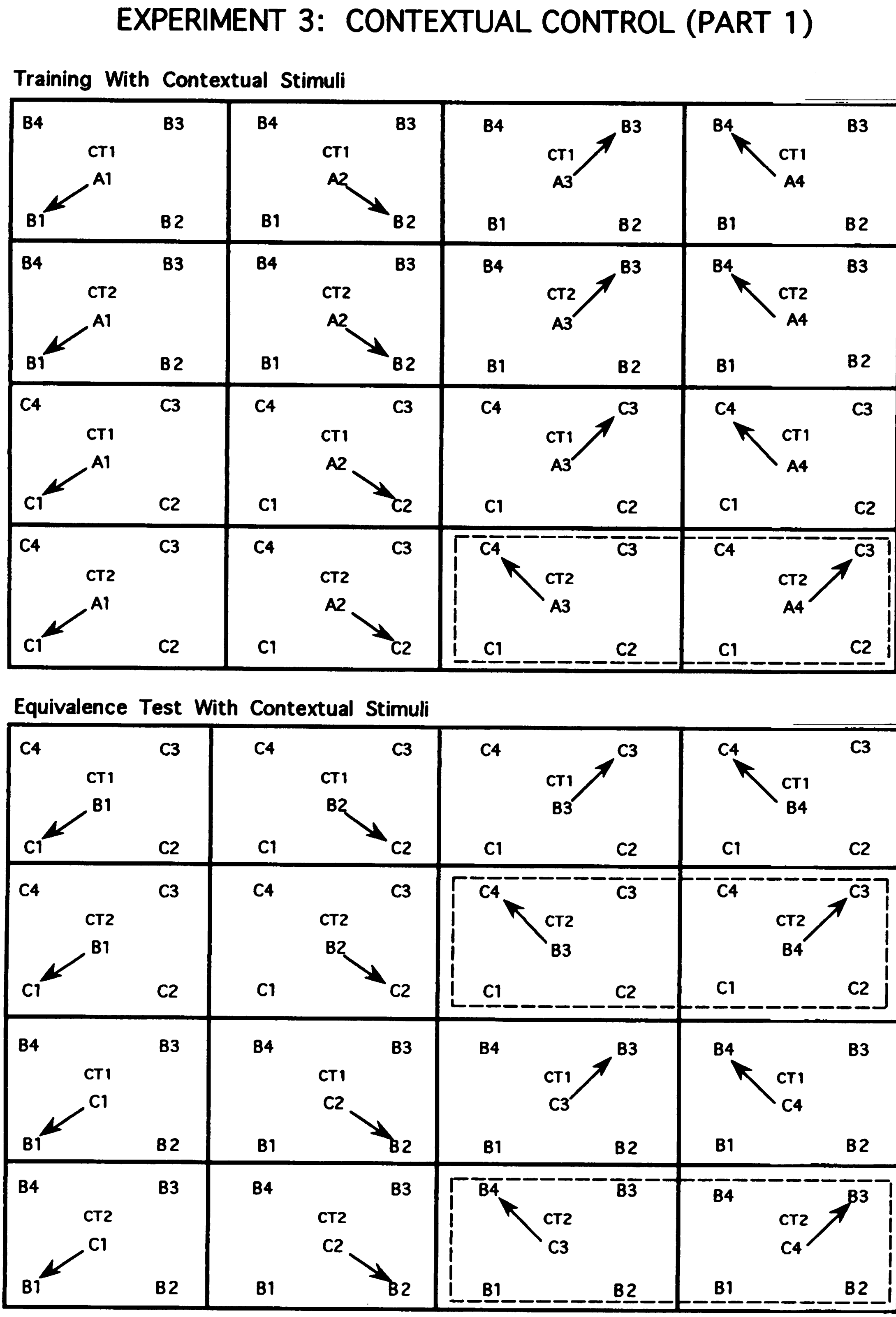


Fig. 7. Schematic representation of the training and testing tasks used in Experiment 3 to establish contextual control over four, three-member equivalence relations. The relations that differed from Experiments 1 and 2 are surrounded by broken lines.

*ulus*, contained two nonequivalent nonsense syllables, subjects would choose the comparison that contained a further two nonequivalent nonsense syllables (e.g., in the presence of Ct1 and given the sample B1C2, subjecs should choose B3C4 rather than B3C3, but in the presence of Ct2, and the same sample, subjects should choose B3C3 rather than B3C4). If a subject produced at least 4 out of 5 correct responses on each of the eight tasks, this performance was defined as contextually controlled equivalence-equivalence responding.

As in Experiments 1 and 2, it was uncertain whether the predicted performances would, in fact, be the most likely outcome, and so a nonpredicted stability criterion was employed for the contextually controlled equivalence-equivalence test. In effect, it was agreed prior to the experiment that a subject's performance would be defined as stable when he or she selected the *same, but not necessarily correct*, comparison stimulus on each of the eight tasks at least four out of five times across a single exposure to one 40 trial block. Furthermore, it was agreed that if a subject did not produce a stable performance on the contextually controlled equivalence-equivalence test, they were to be reexposed to the entire experimental sequence for a second time (i.e., matching-to-sample training, equivalence testing with contextual stimuli, and equivalence-equivalence testing with contextual stimuli). It was also agreed that this recursive training and testing procedure was to be continued until; (i) the subject successfully completed the matching-to-sample training and the contextually controlled equivalence testing, and produced a stable (but not necessarily correct) performance on the equivalence-equivalence test in the presence of the contextual stimuli, or (ii) the subject completed a total of five exposures to the entire experimental sequence, without producing a stable performance on the contextually controlled equivalence-equivalence test. Finally, it was agreed that when either of these two criteria were met, the subject was to be exposed immediately to the Contextual-Stimuli-as-Comparisons-Test.

*Contextual-Stimuli-as-Comparisons-Test*

This test was the same as the equivalence-equivalence test with contextual stimuli, except that the stimuli in each of the eight tasks were reconfigured in the following way; the samples were presented as contextual stimuli, the comparisons were presented as samples, and the contextual stimuli were presented as comparisons (see Figure 8). For example, it was predicted that subjects would choose Ct1 when presented with the sample B3C3 in the presence of B1C1, but choose Ct2 when presented with the sample B3C4 in the presence of B1C1 (i.e., responding in accordance with equivalence-equivalence relations). Similarly, it was predicted that subjects would choose Ct1 when presented with the sample B1C2 in the presence of B3C4, but choose Context 2 when presented with the sample B1C2 in the presence of B3C3 (i.e., responding in accordance with nonequivalence-nonequivalence relations).

As indicated above, the eight tasks were presented five times each, in a quasi-random order, in 40 trial blocks without feedback. It was again uncertain whether the predicted performances would be the most likely outcome, and so the nonpredicted stability criterion was also employed for the contextual-stimuli-as-comparisons-test (i.e., choose the same, but not necessarily correct, comparison on each task at least four out of five times across a single exposure to one 40 trial block). It was agreed prior to the experiment, that if a subject did not produce a stable performance on the contextual-stimuli-as-comparisons-test, he or she would be re-exposed to the entire experimental sequence for a second time (i.e., matching-to-sample training, equivalence testing with contextual stimuli, equivalence-equivalence testing with contextual stimuli, and testing with contextual-stimuli-as-comparisons). It was also agreed that this recursive training and testing procedure would be continued until; (i) the subject successfully completed the matching-to-sample training and the contextually controlled equivalence testing, and produced a stable (but not necessarily correct) performance on the equivalence-

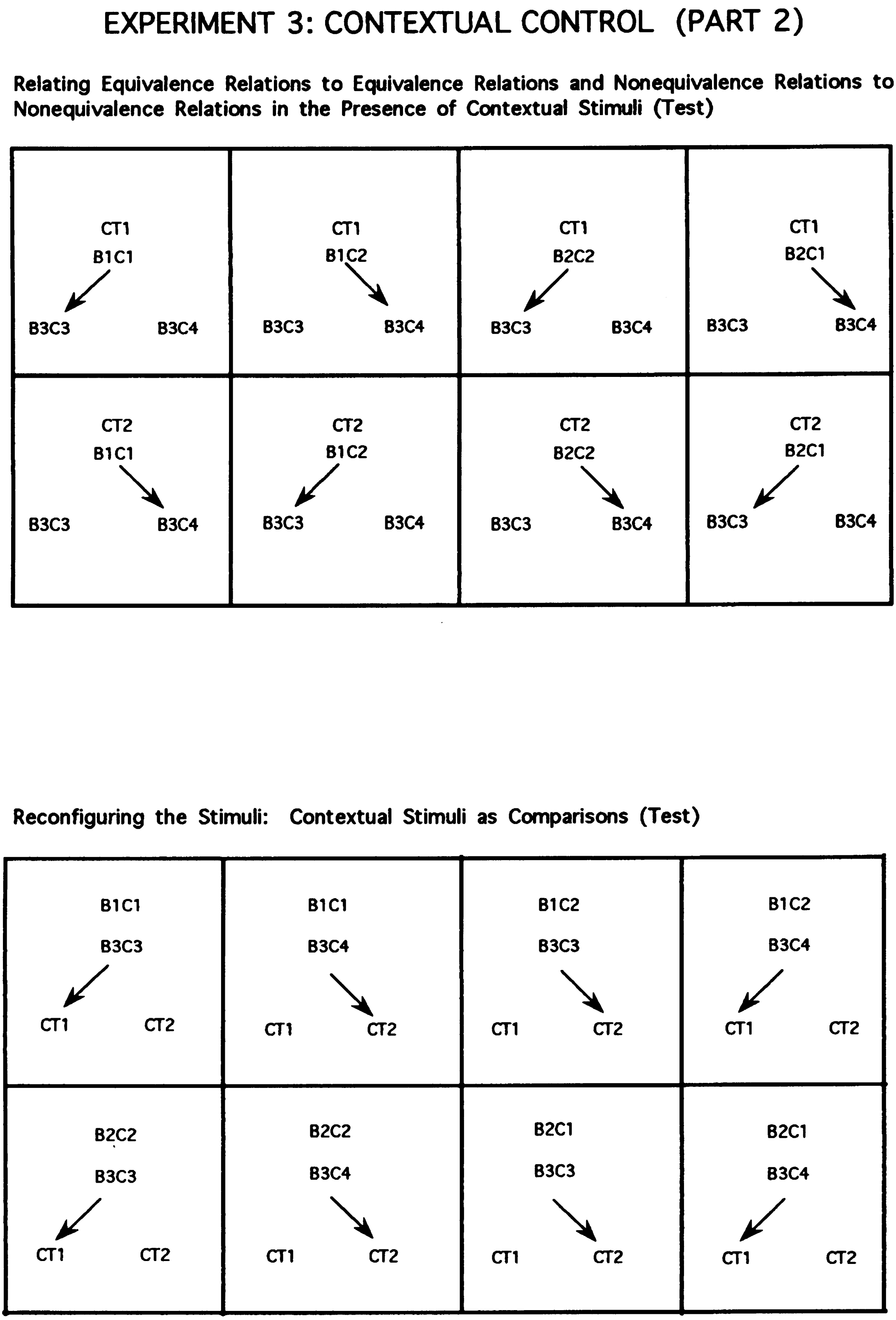


Fig. 8. Schematic representation of the tasks used in Experiment 3 to test for contextual control over equivalence-equivalence responding (upper panel) and to test for the maintenance of contextual control when the stimuli in the testing tasks were reconfigured (lower panel).

equivalence test in the presence of the contextual stimuli, and produced a stable (but not necessarily correct) performance on the contextual-stimuli-as-comparisons-test, or (ii) the subject completed five exposures to the entire experimental sequence, without producing a stable performance on the contextually controlled equivalence-equivalence test and/or without producing a stable performance on the contextual-stimuli-as-comparisons-test.

## Results and Discussion

Figure 9 presents the total number of training trials per exposure, the percentage of correct responses made during each exposure to the contextually controlled equivalence test, to the contextually controlled equivalence-equivalence test, and to the contextual-stimuli-as-comparisons-test (see appendix for a detailed breakdown of each subject's performance).

The data for the experimentally naive subject (Subject 1), will be described first and in the most detail. On his first exposure, this subject required a total of 192 training trials. He then failed the contextually controlled equivalence test (27.5 percent correct), and produced an unstable performance on the contextually controlled equivalence-equivalence test (47.5 percent correct). After a further 88 training trials, he again failed the contextually controlled equivalence test although his performance improved substantially (87.5 percent correct), and he produced another unstable (but somewhat improved) performance on the contextually controlled equivalence-equivalence test (77.5 percent correct). The subject was retrained (32 training trials), and was successfully retested for contextually controlled equivalence responding (100 percent correct), and for contextually controlled equivalence-equivalence responding (100 percent correct). He was then reexposed to the contextual-stimuli-as- comparisons-test, and he produced a stable and 97.5 percent correct performance.

The remaining two subjects from Experiment 3 also completed the matching-to-sample training, demonstrated

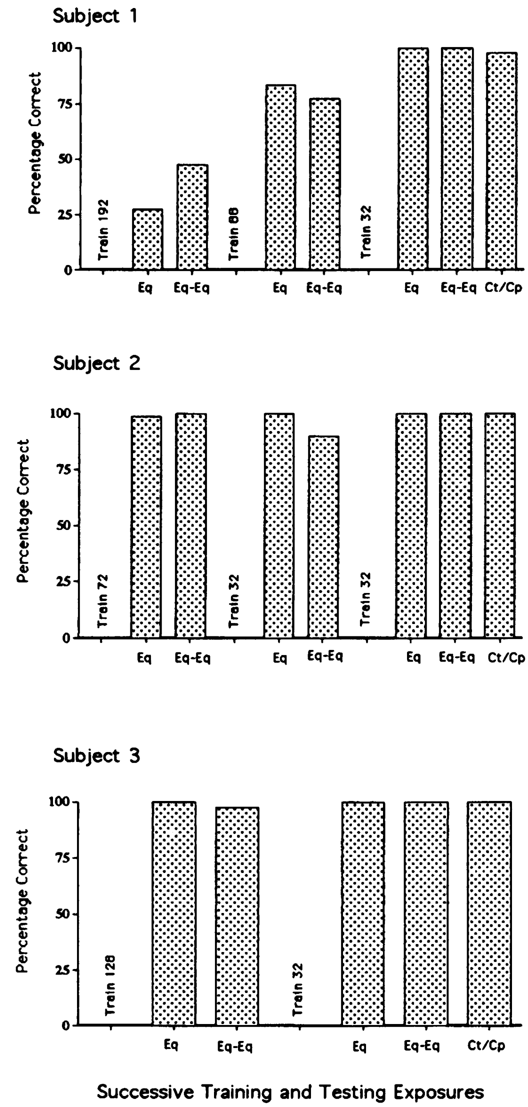


Fig. 9. Percentage of correct responses across successive exposures to the training and testing stages of Experiment 3 for Subjects 1 to 3. Eq indicates a test for contextually controlled equivalence responding, Eq-Eq indicates a test for contextually controlled equivalence-equivalence responding, and Ct/Cp indicates a test for contextually controlled equivalence-equivalence responding when the contextual cues were presented as comparisons.

contextually controlled equivalence responding, contextually controlled equivalence-equivalence responding, and produced the predicted responses when the contextual stimuli were used as comparisons. Subject 2 required a total of 136 training trials, and was provided with three exposures to the equivalence test, three exposures to the equivalence-equivalence test, and one exposure to the final test with contextual stimuli used as comparisons. Subject 3 required a total of 160 training trials, two exposures to the equivalence test, two exposures to the equivalence-equivalence test, and one exposure to the test with contextual stimuli as comparisons.

The data from Experiment 3 clearly showed that it was possible to demonstrate contextual control over equivalence relations and equivalence-equivalence relations, and that subjects' contextually controlled equivalence-equivalence responding can be maintained when the samples are presented as contextual stimuli, the comparisons are presented as samples, and the contextual stimuli are presented as comparisons. One experimentally naive (non-graduate) subject, and two non-naive (graduate) subjects demonstrated these behavioral effects.

## GENERAL DISCUSSION

The data from Experiment 1 showed that a range of subjects, including a 12 year old boy, could successfully relate equivalence relations to other, separate equivalence relations and could relate nonequivalence relations to other, separate nonequivalence relations, in the absence of explicit reinforcement. In Experiment 2, the same procedures were employed as in Experiment 1, except that exposure to the equivalence-equivalence test came before exposure to the standard equivalence test. The results allowed us to conclude that exposure to the standard equivalence test was not a prerequisite for passing the equivalence-equivalence test. In Experiment 3, the data clearly showed that it was possible to demonstrate contextual control over equivalence relations and equivalence-equivalence relations. Furthermore, it was found that subjects' contextually controlled equivalence-equivalence responding was maintained when the samples were presented as contextual stimuli, the comparisons were presented as samples, and the contextual stimuli were presented as comparisons.

It is important to note that: (i) all three experiments employed a pre-determined stability criterion during the equivalence-equivalence testing (i.e., stable but not necessarily correct), and (ii) all experimental subjects throughout the study showed equivalence-equivalence responding. It is very likely, therefore, that the predicted performances were largely derived from the trained relations and not from the additional feedback provided by repeated training and testing that is often employed in equivalence research (see Barnes & Keenan, 1993, p.63).

The fact that two children (a 12 year old in Experiment 1, and a 9 year old Experiment 2) demonstrated equivalence-equivalence responding indicates that this form of complex stimulus control does not necessarily require an extended educational training typically acquired during second and third level schooling. This finding suggests, therefore, that equivalence-equivalence responding is not a by-product of advanced logical or mathematical reasoning; in fact, it might be argued that equivalence-equivalence responding, as an instance of relational framing, constitutes a functional-analytic interpretation of some types of reasoning abilities (see Hayes, 1994). Given this possibility, a future direction for studies of equivalence-equivalence responding might be to conduct a cross-sectional, developmental study with children of different ages. Such a study would allow us to determine whether preschoolers are capable of equivalence-equivalence responding, and if not, exactly when this performance emerges in the behavior of the developing child. Furthermore, because no children were employed in Experiment 3 of the current study, a future experiment might attempt to demonstrate contextual control over equivalence-equivalence responding in children of various ages.

Future research might also examine the verbal utterances made by subjects as they interact with the training and testing tasks. This could be achieved using a think aloud procedure and protocol analysis (Hayes, 1986). Of course, these data would not explain the equivalence-equivalence responding (Barnes, 1989; Hayes &

Brownstein, 1986), but they may provide some important clues for tracking down the behavioral histories responsible for such performances. For example, the subjects' verbal utterances may have functioned as "prompts" for responding, and thus it might be useful to ask what variables were responsible for these utterances, and how they came to "prompt" or control subjects' responding. Perhaps some of the utterances could be analyzed as tacts (Skinner, 1957) evoked by the experimental stimuli combined with intraverbals (Skinner, 1957) evoked by the instructions. A history of autoclitic (Skinner, 1957) behavior may have then caused the tacts and intraverbals to acquire their "prompting" or controlling properties over the subjects' matching-to-sample performances. In conducting this work, however, we would need to distinguish between tacts, intraverbals, and autoclitics that have been explicitly reinforced, and those that are derived in some sense. In fact, Skinner (1957) himself coined the phrase *autoclitic frame* to describe autoclitics that combine already established behavioral units into novel utterances. In short, we believe the distinction between derived and non-derived relational responding to be an important one, and making this distinction clear will be a critical step towards a complete functional analysis of human language.

In Experiments 1 and 2, the equivalence-equivalence test was presented in two blocks of 20 trials. The first block presented samples that contained stimuli from the equivalence relations numbered 1 and 2 and comparisons that contained stimuli from equivalence relations numbered 3 and 4 (e.g., B1C1 as sample with B3C3 and B3C4 as the "correct" and "incorrect" comparisons, respectively). In effect, none of the individual elements in the sample stimuli were present in either the "correct" or "incorrect" comparison stimuli (see Figure 2). In contrast, the second block of test trials presented samples and "incorrect" comparisons that both contained an element in common (e.g., *B1*C1 as sample with B4C4 and *B1*C2 as the "correct" and "incorrect" comparisons, respectively). When first designing the current experiments, we thought that the second block of equivalence-equivalence test trials might produce a greater number of errors than the first block. Our reasoning was as follows. Insofar as reflexivity is a defining feature of equivalence responding, then presenting a sample and an incorrect comparison that both contain an element in common might generate competition between responding in accordance with arbitrary equivalence relations and responding in accordance with the non-arbitrary relation of reflexivity. In the event, however, the subjects' final performances on the second block of the equivalence-equivalence test were similar to their performances on the first block (see appendix). This finding therefore supports the conclusion that subjects were selecting a comparison that was *arbitrarily* equivalent to the sample, rather than physically similar to the sample stimulus. Perhaps future studies might utilise this strategy (i.e., presenting complex samples and comparisons with elements in common) to further analyze responding in accordance with the arbitrary relations of symmetry and transitivity, and the non-arbitrary relation of reflexivity (see Hayes, 1991, p. 32; Saunders & Green, 1992, p. 236; Sidman, 1994, p. 167).

Although the final performances of the subjects on the first and second blocks of the equivalence-equivalence test did not differ, it is interesting that during their first exposure to the equivalence-equivalence test, Subjects 1 and 2 (Experiment 1) both failed to pass the first test block, but successfully passed the second (see appendix for performance breakdown). These "acquisition data" could be taken to indicate that the tasks used in the second block of equivalence-equivalence test trials provided, in some undefined way, a more effective context for equivalence responding than the first block. These data (from Subjects 1 and 2) appear, therefore, to contradict our original expectations. Nevertheless, it is important to note that the improvement in the second block may have resulted from a simple order effect, and thus future researchers would be well advised to counterbalance the order of presentation of the "first" and "second" blocks of the equivalence-equivalence test.

The subjects in Experiment 3 all produced the predicted patterns of responding when the contextual stimuli were presented as comparisons. These patterns appeared to be consistent with the contextual control that had previously been demonstrated over the equivalence and equivalence-equivalence relations. For example, when subjects were presented with B1C1 as a contextual stimulus and B3C3 as a sample, all three subjects reliably picked Ct1. In effect, the subjects appeared to respond to the previously established contextual functions by choosing Ct1 as a comparison, *because previously in the presence of Ct1*, the two elements of the new contextual stimulus (B1C1) were from the same equivalence class, and the two elements from the new sample stimulus (B3C3) were from the same class. Thus, when subjects chose Ct1 it was still functioning as a contextual stimulus, for the equivalence-equivalence relation between B1C1 and B3C3, but was also functioning *simultaneously* as a comparison within the matching-to-sample task. Interestingly, a number of researchers have suggested that contextual stimuli may enter into the matching-to-sample task itself (Markham & Dougher, 1993; Sidman, 1994, p. 514-528; Stromer & Mackay, 1992; Stromer, McIlvane, & Serna, 1993; Stromer & Stromer, 1990a, 1990b; see Barnes, 1994, for a discussion of how the respondent-type functions of sample and comparison stimuli may function as contextual cues for equivalence responding). The current data support this idea, and moreover they indicate considerable flexibility in the types of emergent performances that can emerge given appropriate testing conditions. Perhaps, future research in this area might attempt to discover whether there is a specific limit to the ways in which equivalence testing tasks can be reconfigured before previously established derived relational responding fails to transfer to the reconfigured tasks.

The procedures used in this study were reasonably effective in obtaining the predicted performances, but some improvements might be made. For example, in Experiment 1 only Subject 5 successfully passed the standard equivalence test on his first exposure (although Subjects 4 and 6 both failed by only two 2 incorrect responses). Interestingly, recent research has found that training and testing for symmetry relations before training and testing for more complex relations (e.g., combined symmetry and transitivity) appears to facilitate accurate responding in accordance with equivalence relations (Fields, Adams, Newman, & Verhave, 1992). Within the context of the current study, therefore, the failure by most subjects to pass the equivalence test on their first exposure may be related to the fact that they were not tested for symmetry responding before being tested for the formation of equivalence relations. Furthermore, it is possible that passing the equivalence-equivalence test also might have been facilitated in Experiment 1, if subjects had previously been tested for symmetry relations. Future research could certainly examine this issue.

An interesting feature of the current study emerges when one compares the data from Experiment 1 with that of Experiment 2. Consider the following. Subjects in Experiment 2, having passed the equivalence-equivalence test, immediately passed the standard equivalence test (i.e., on their first exposure). This contrasts with Experiment 1, in which five of the six subjects failed the equivalence-equivalence test on their first exposures, although they had successfully passed the standard equivalence test. Why did the equivalence-equivalence test appear to generate the equivalence relations and the equivalence-equivalence relations, but the standard equivalence test generate only the equivalence relations? One possible reason may be the different formats used in the equivalence and equivalence-equivalence tests; the equivalence-equivalence test presented compound stimuli and only two comparisons, whereas the baseline training and standard test used four comparisons and single nonsense syllables. Thus, when the equivalence-equivalence test was first introduced, this change in format may have initially disrupted a transfer of the previously established equivalence responding to the new equivalence-equivalence test. This failure (due to a novel format) would have been less likely in Experiment 2 (when shifting from the equivalence-equivalence test to the equivalence test) because subjects had already been exposed to the standard format during the baseline conditional discrimination training. A future study might therefore attempt to replicate Experiment 1, but also provide the subjects with a pretraining history in which they are exposed to the format used in the equivalence-equivalence test (i.e., a compound sample, and two compound comparisons). If subjects exposed to such a pretraining history pass the equivalence-equivalence test immediately after passing the standard equivalence test, this would indicate, contrary to the present findings, that responding in accordance with standard equivalence relations normally provides a sufficient basis for responding in accordance with equivalence-equivalence relations.

*Analogical reasoning and cognition*

In the introduction we suggested that the study of equivalence-equivalence relations may represent an appropriate starting point for a behaviour-analytic interpretation of analogical reasoning. In spite of the fact that cognitive psychologists have examined analogical reasoning from a variety of perspectives, it has been argued that the research fails to explicate the underlying causes of analogical reasoning (Ortony, Reynolds, & Arter, 1978). Cognitive psychologists have defined analogical reasoning, for example, as a transfer of relational knowledge from one domain to another. This process is divided into several components, the most common ones being: accessing the source, mapping between the base and the target domain, and the production of more general schemata. Thus, analogical reasoning is defined in terms such as: "accessing" and "mapping" which in themselves lack clear definition (Ortony, et al., 1978). In addition many cognitive theories of analogical reasoning are task-based and are not part of a larger theory. The data from the present study, however, indicate that it should be possible to analyze analogical reasoning using the behavior-analytic framework of relational frame theory, and that this analysis may be used to include other factors such as contextual control over analogical reasoning (i.e., contextual control over equivalence and equivalence-equivalence responding).

The current study may also have implications for a theoretical analysis of important aspects of cognition such as language and intelligence. Although "Differential psychologies have long recognised the close relation between analogical reasoning and intelligence" (Sternberg, 1977, p. 353), traditional definitions of intelligence have lacked an understanding of the flexibility and development of the underlying behavioural processes involved in performances typically categorized as intelligent. Recently, however, Hayes (1994) has argued that it should be possible to train equivalence and other relational activities as operant behavior, and the subsequent improvement in relational responding should lead to improved abilities in areas of cognition, such as language and intelligence. This strategy avoids the typical approach to language and intelligence taken by cognitive psychology which has tended to emphasize "content" by the training of specific words and/or the acquisition of specific concepts applicable in the real world. While these are important, the theoretical implications of the relational frame approach suggest that the flexibility and development of the underlying behavioral processes are equally important. For example, consider a classroom setting where games could be designed to improve the flexibility of a child's relational responding. Questions could be asked such as: "If x is the same as y, and y is the same as z, do I like z if I like x?" Or for younger children, the games could be simplified to include questions such as: "If the teddy-bear is called Fluffy please bring Fluffy over here." In addition, an examination of analogical reasoning, using relational frame theory, could help provide a functional-analytic definition of intelligence in terms of the degree of competence shown in complex testing tasks. In any case, the current research clearly indicates that behavior analysis can provide a unique and productive approach to an area of human behavior that is typically viewed as the sole concern of cognitive psychology.

## REFERENCES

Barnes, D. (1989). Behavior-behavior analysis, human schedule performance, and radical behaviorism. *The Psychological Record, 39,* 339-350.

Barnes, D. (1994). Stimulus equivalence and relational frame theory. The *Psychological Record, 44,* 91-124.

Barnes, D., Browne, M., Smeets, P.M., & Roche, B. (1988). A transfer of functions and a conditional transfer of functions through equivalence relations in three to six year old children. *The Psychological Record, 45,* 405-430.

Barnes, D., & Holmes, Y. (1991). Radical behaviorism, stimulus equivalence, and human cognition. *The Psychological Record, 41,* 19-31.

Barnes, D., & Roche, B. (1996). Relational frame theory and stimulus equivalence are fundamentally different: A reply to Saunders' commentary. *The Psychological Record, 46,* 489-507.

Bush, K.M., Sidman, M., & deRose, T. (1989). Contextual control of emergent equivalence relations. *Journal of the Experimental Analysis of Behavior, 51,* 29-45.

Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination functions in accordance with the arbitrarily applicable relations of sameness, more than, and less than. *Journal of the Experimental Analysis of Behavior, 64,* 163-184.

Fields, L., Adams, B.J., Newman, S., & Verhave, T. (1992). Interactions among emergent relations during equivalence class formation. *The Quarterly Journal of Experimental Psychology, 45 B,* 125-138.

Fields, L., Adams, B.J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53,* 345-358.

Gick, M.L., & Holyoak, K.J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1-38.

Hayes, S.C. (1991). A relational control theory of stimulus equivalence. In L.J. Hayes & P.N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 109-118). Reno, NV: Context Press.

Hayes, S.C. (1994). Relational frame theory: A functional approach to verbal events. In S.C. Hayes, L.J. Hayes, M. Sato, & K. Ono (Eds.). *Behavior analysis of language and cognition* (pp. 11-30). Reno NV: Context Press.

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

Hayes, S. C., & Brownstein, A. J. (1986). Mentalism, behavior-behavior relations, and the behavior-analytic view of the purposes of science. *The Behavior Analyst, 9,* 175-190.

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.

Ortony, A., Reynolds, R.E., & Arter, J.A. (1978). Metaphor: Theoretical and empirical research. *Psychological Bulletin, 85,* 919-943.

Perez-Gonzalez, L.A. (1994). Transfer of relational stimulus control in conditional discriminations. *Journal of the Experimental Analysis of Behavior, 61,* 487-503.

Saunders, R.R., & Green, G. (1992). The nonequivalence of behavioral and mathematical equivalence. *Journal of the Experimental Analysis of Behavior, 57,* 227-241.

Sidman, M. (1992). Equivalence relations: Some basic considerations. In S.C. Hayes & L.J. Hayes (Eds.), *Understanding verbal relations* (pp. 15-27). Reno, NV: Context Press.

Sidman, M. (1994). *Equivalence relations: A research story.* Boston, MA: Authors Cooperative.

Skinner, B. F. (1957). *Verbal behavior.* New York Appleton-Century-Crofts.

Sternberg, R.J. (1977). Component processes in analogical reasoning. *Psychological Review, 84,* 353-378.

Stromer, R., & Mackay, H.A. (1992). Spelling and emergent picture-printed word relations established with delayed identity matching to complex samples. *Journal of Applied Behavior Analysis, 25,* 893-904.

Stromer, R., McIlvane, W.J., & Serna, R.W. (1993). Complex stimulus control and equivalence. *The Psychological Record, 43,* 585-598.

Stromer, R., & Stromer, J.B. (1990a). The formation of arbitrary stimulus classes in matching to complex samples. *The Psychological Record, 40,* 51-66.

Stromer, R., & Stromer, J.B. (1990b). Matching to complex samples: Further study of arbitrary stimulus classes. *The Psychological Record, 40,* 505-516.

Wulfert, E., Greenway, D.E., & Dougher, M.J., (1994). Third order equivalence classes. *The Psychological Record, 44,* 411-440.

Wulfert, E., & Hayes, S.C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50,* 25-141.

## APPENDIX 1

Number of test trials on which responses were consistent with the trained relations in Experiment 1.

| | | | EQUIVALENCE RELATIONS | | | | | | | | EQUIVALENCE-EQUIVALENCE RELATIONS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sub# | Training trials | Exposure# | B1C1 | B2C2 | B3C3 | B4C4 | C1B1 | C2B2 | C3B3 | C4B4 | B1C1 B3C3 | B1C2 B3C4 | B2C2 B3C3 | B2C1 B3C4 | B1C1 B4C4 | B4C3 B1C2 | B2C2 B4C4 | B3C4 B1C2 |
| S1 | 184 | 1 | 2 | 0 | 2 | 0 | 0 | 3 | 1 | 1 | | | | | | | | |
| | 208 | 2 | 4 | 5 | 5 | 4 | 5 | 3 | 4 | 5 | | | | | | | | |
| | 60 | 3 | 5 | 5 | 5 | 5 | 5 | 4 | 3 | 5 | | | | | | | | |
| | 20 | 4 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 1 | 2 | 0 | 0 | 5 | 5 | 5 | 5 |
| | 20 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 4 | 4 | 4 | 5 | 5 | 5 | 4 |
| S2 | 132 | 1 | 5 | 3 | 1 | 0 | 4 | 2 | 0 | 0 | | | | | | | | |
| | 168 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | | | | | | | | |
| | 20 | 3 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 3 | 5 | 5 | 5 | 5 |
| | 20 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 |
| | 20 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| S3 | 68 | 1 | 3 | 2 | 1 | 2 | 3 | 5 | 0 | 1 | | | | | | | | |
| | 24 | 2 | 4 | 5 | 0 | 0 | 5 | 5 | 0 | 1 | | | | | | | | |
| | 20 | 3 | 1 | 1 | 2 | 3 | 4 | 0 | 0 | 4 | | | | | | | | |
| | 28 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | | | | | | | | |
| | 20 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 |
| | 20 | 6 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| S4 | 172 | 1 | 3 | 4 | 3 | 4 | 5 | 5 | 5 | 5 | | | | | | | | |
| | 36 | 2 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | | | | | | | | |
| | 20 | 3 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 4 | 3 | 4 | 3 | 4 | 1 | 5 | 1 |
| | 36 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| S5 | 44 | 1 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 1 | 5 | 2 | 5 | 5 | 5 | 3 |
| | 20 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| | 20 | 3 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| S6 | 180 | 1 | 5 | 2 | 4 | 5 | 5 | 4 | 5 | 5 | 2 | 3 | 0 | 1 | 1 | 1 | 0 | 0 |
| | 36 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 1 | 5 | 5 | 5 | 2 | 5 |
| | 32 | 3 | 5 | 4 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| | 20 | 4 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |

## APPENDIX 2

Number of test trials on which responses were consistent with the trained relations in Experiment 2.

| Sub# | Training trials | Exposure# | EQUIVALENCE-EQUIVALENCE RELATIONS B1C1 B3C3 | B1C2 B3C4 | B2C2 B3C3 | B2C1 B3C4 | B1C1 B4C4 | B4C3 B1C2 | B2C2 B4C4 | B3C4 B1C2 | EQUIVALENCE RELATIONS B1C1 | B2C2 | B3C3 | B4C4 | C1B1 | C2B2 | C3B3 | C4B4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S1 | 136 | 1 | 4 | 2 | 0 | 3 | 2 | 2 | 4 | 1 | | | | | | | | |
| | 60 | 2 | 1 | 4 | 1 | 5 | 3 | 2 | 2 | 4 | | | | | | | | |
| | 40 | 3 | 2 | 5 | 3 | 2 | 5 | 3 | 3 | 4 | | | | | | | | |
| | 20 | 4 | 5 | 0 | 1 | 3 | 1 | 5 | 5 | 1 | | | | | | | | |
| | 36 | 5 | 5 | 5 | 5 | 4 | 4 | 2 | 4 | 2 | | | | | | | | |
| | 24 | 6 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| S2 | 248 | 1 | 3 | 1 | 4 | 3 | 0 | 0 | 0 | 0 | | | | | | | | |
| | 24 | 2 | 4 | 0 | 5 | 1 | 5 | 3 | 1 | 4 | | | | | | | | |
| | 36 | 3 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | | | | | | | | |
| | 20 | 4 | 5 | 5 | 3 | 5 | 5 | 5 | 5 | 5 | | | | | | | | |
| | 20 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| S3 | 276 | 1 | 0 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | | | | | | | | |
| | 96 | 2 | 2 | 2 | 3 | 1 | 1 | 2 | 4 | 3 | | | | | | | | |
| | 20 | 3 | 4 | 5 | 4 | 3 | 5 | 5 | 5 | 3 | | | | | | | | |
| | 36 | 4 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| S4 | 88 | 1 | 2 | 0 | 4 | 5 | 5 | 5 | 0 | 0 | | | | | | | | |
| | 36 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | | | | | | | | |
| | 20 | 3 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 |
| S5 | 164 | 1 | 0 | 1 | 1 | 2 | 3 | 3 | 5 | 1 | | | | | | | | |
| | 56 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | | | | | | | | |
| | 28 | 3 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 4 | 5 | 5 |
| S6 | 120 (approx) | 1 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | | | | | | | | |
| | 20 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |

## APPENDIX 3

Number of test trials on which equivalence and equivalence-equivalence responses were consistent with the trained relations in Experiment 3.

| | | | EQUIVALENCE RELATIONS | | | | | | | | | | | | EQUIVALENCE-EQUIVALENCE RELATIONS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Ctx Cue 1 | | | | Ctx Cue 2 | | Ctx Cue 1 | | | | Ctx Cue 2 | | Ctx Cue 1 | | | | Ctx Cue 2 | | | |
| Sub# | Training trials | Exposure# | B1C1 | B2C2 | B3C3 | B4C4 | B3C4 | B4C3 | C1B1 | C2B2 | C3B3 | C4B4 | C3B4 | C4B3 | B1C1 B3C3 | B1C2 B3C4 | B2C2 B3C3 | B2C1 B3C4 | B1C1 B3C4 | B1C2 B3C3 | B2C2 B3C4 | B2C1 B3C3 |
| S1 | 192 | 1 | 3 | 0 | 1 | 2 | 0 | 1 | 0 | 3 | 2 | 3 | 2 | 1 | 0 | 1 | 4 | 4 | 2 | 1 | 3 | 4 |
| | 88 | 2 | 5 | 3 | 4 | 4 | 2 | 4 | 5 | 5 | 4 | 4 | 5 | 4 | 3 | 3 | 5 | 4 | 5 | 4 | 3 | 4 |
| | 32 | 3 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| S2 | 72 | 1 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| | 32 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 5 |
| | 32 | 3 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| S3 | 128 | 1 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 |
| | 32 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |

## APPENDIX 4

Number of test trials on which responses on the Contextual-Stimuli-as-Comparisons-Test were consistent with the trained relations in Experiment 3.

| | CONTEXTUAL CUE AS COMPARISON | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | B1C1 B3C3 Ct1 | B1C1 B3C4 Ct2 | B1C2 B3C3 Ct1 | B1C2 B3C4 Ct1 | B2C2 B3C3 Ct1 | B2C2 B3C4 Ct2 | B2C1 B3C3 Ct2 | B2C1 B3C4 Ct1 |
| S1 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 |
| S2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| S3 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |