

RESPONSE LATENCIES TO MULTIPLE
DERIVED STIMULUS RELATIONS: TESTING
TWO PREDICTIONS OF RELATIONAL FRAME THEORY

DENIS O'HORA, BRYAN ROCHE, and DERMOT BARNES-HOLMES
National University of Ireland, Maynooth

PAUL M. SMEETS
Leiden University, Holland

In Experiment 1, 3 college students were exposed to relational pretraining to establish the contextual functions of Same, Opposite, More Than, and Less Than in four arbitrary stimuli. Subjects were then trained on the matching-to-sample tasks A1-B1 and Y1-N1, in the presence of the More-Than contextual cue, A1-B2 and Y1-N2 in the presence of the Less-Than contextual cue, C1-D1 and E1-D2 in the presence of the Same cue, and C1-D2 and E1-D1 in the presence of the Opposite cue. Test trials were subsequently administered to probe for the mutually entailed relations; Less-Than/B1-A1, Less-Than/N1-Y1, More-Than/B2-A1, More-Than/N2-Y1, Same/D1-C1, Same/D2-E1, Opposite/D2-C1, and Opposite/D1-E1. Response latencies to probes for derived Same/Opposite relations were significantly lower than those for derived More Than/Less Than relations. Experiment 2 exposed 4 subjects to training across each of the four relations and used a novel stimulus set to test for reduced response latencies to the derived relations. Response latencies to More-Than/Less-Than probes reduced significantly across the original to the novel stimulus set, whereas latencies to Same/Opposite probes were low across both stimulus sets.

When human subjects are trained on a series of conditional discriminations, the stimuli involved may control responding in ways that are not readily predicted using traditional behavioral principles. For example, if choosing the arbitrary stimuli, B1 and C1, is reinforced on separate trials, given a further stimulus, A1, then, the presentation of B1 or C1 may reliably occasion the choosing of A1 (symmetry) without reinforcement. Furthermore, the presentation of B1 may control choosing C1, and the presentation of C1 may control choosing B1 (combined symmetry and transitivity), again without reinforcement. These effects are collectively referred to as stimulus equivalence and the stimuli A1, B1, and C1 are said to participate in an equivalence relation (Barnes, 1994; Barnes & Holmes, 1991; Hayes, 1991; Sidman, 1971, 1990).

Requests for reprints should be addressed to Bryan Roche, Department of Psychology, National University of Ireland, Maynooth, Co. Kildare, Ireland. The authors thank Michael Clayton and an anonymous reviewer for their helpful comments. The first three authors say farewell to Cork, where so many ideas were born.

The possibility that human subjects can respond in accordance with relations other than equivalence was examined in a study by Steele and Hayes (1991). Subjects were trained to relate Same stimuli (e.g., a large line with a large line) in the presence of one contextual cue, Opposite stimuli (e.g., a large line with a small line) in the presence of a second contextual cue, and Different stimuli (e.g., a circle with a cross) in the presence of a third contextual cue. Subsequently, subjects were trained in an extensive network of conditional discriminations, with each discrimination being made in the presence of one of the three contextual cues. To understand the procedures involved in this complex experiment, consider the following six training trials; Same/A1/B1-B2-B3, Same/A1/C1-C2-C3, Opposite/A1/B1-B2-B3, Opposite/A1/C1-C2-C3, Different/A1/B1-B2, Different/A1/C1-C2, where A1 represents the sample, and the B and C stimuli represent the comparisons. Choosing B1 and C1 was always reinforced in the presence of the Same stimulus, choosing the B3 and C3 stimuli was always reinforced in the presence of the Opposite stimulus, and choosing the B2 and C2 stimuli was always reinforced in the presence of the Different stimulus.

Now consider the following three test trials; Same/B1/C1-C2-C3, Same/B3/C1-C2-C3, Different/C1/B1-B2-N3. Subjects chose C1, C3, and B2, respectively on these tasks, indicating that the relations of Same, Opposite, and Different had been derived. More specifically, if B1 and C1 are the Same as A1, then B1 and C1 are the Same, and if B3 and C3 are Opposite to A1 then B3 and C3 are the Same. Finally, if B2 is Different from A1, and C1 is the Same as A1, then B2 is also Different from C1. Several further experimental studies have demonstrated that human subjects can be trained to respond to a variety of derived stimulus relations, including; Same and Opposite (Dymond & Barnes, 1996; Roche & Barnes, 1996, 1997; Roche, Barnes-Holmes, Smeets, Barnes-Holmes, & McGeady, 2000), Different (Roche & Barnes, 1996), and More-Than and Less-Than (Dymond & Barnes, 1995; see also Barnes-Holmes, Barnes-Holmes, & Cullinan, 2001).

All of the foregoing studies of derived stimulus relations were conducted under the rubric of Relational Frame Theory (RFT). This theory explicitly recognizes and attempts to explain the emergence of a variety of derived stimulus relations other than equivalence. Given that the current research was conducted as part of the ongoing relational frame research program at our laboratory, we should briefly outline the essence of the relational frame approach.

Explaining Additional Stimulus Relations: Relational Frame Theory

Both humans and animals can respond to a wide variety of nonarbitrary stimulus relations (e.g. Reese, 1968). Relational Frame Theory argues, however, that at least some organisms can learn to respond in accordance with arbitrary stimulus relations. More specifically, given a sufficient history of training in nonarbitrary relational responding (e.g., discriminating stimuli based on physical magnitude), these relations

may be applied arbitrarily to any novel set of stimuli in an appropriate context (e.g., discriminating stimuli based on monetary “magnitude”). In effect, RFT argues that nonarbitrary relational responding can become generalized such that it is arbitrarily applicable to any set of stimuli. From this perspective, derived relational responding is a form of generalized operant behavior (see Healy, Barnes, & Smeets, 1998, for empirical evidence to this effect).

As an example of the foregoing, consider an individual with an extensive personal history of nonarbitrary relational responding to the relations of More Than and Less Than. During the early phases of this training (e.g., by parents and teachers in naturalistic settings) the person is trained to discriminate a wide variety of objects (e.g., food, toys, etc.) on the basis of physical size. Given sufficiently extensive and explicit reinforcement for responding correctly to this nonarbitrary relation, responding may generalize such that the child can reliably discriminate the bigger of any two objects. With continued training of this type, across multiple exemplars, the relational responding repertoire will generalize further still such that it becomes applicable to events that are unrelated along any physical continuum. For instance, in the context of size, a person with an appropriate history of responding to nonarbitrary “greater than” relations, will respond to a nickel as “greater than” a dime (i.e., based on physical size). However, when the comparison relation becomes arbitrarily applicable (e.g., in the context of value) this person may respond to a dime as greater than a nickel (i.e., a dime is of greater *arbitrary* value). These types of arbitrary relational responses are controlled by context (e.g., value) rather than by formal properties of the stimuli being related. In this sense, such responses are arbitrarily applicable and the relations that define this application are called relational frames (Hayes, 1991, Hayes & Hayes, 1989).

The relational frames of particular interest in the current study are those of coordination, opposition, and comparison. Coordination is the relation of identity, sameness or similarity (“this is that”). This relation likely has the most protracted use within the verbal community (Hayes & Hayes, 1989; Roche & Barnes, 1996). Indeed, much of the earliest language training received by children seems to be heavily dependent upon it. The frame of coordination, therefore, is probably the first to be abstracted sufficiently that it becomes arbitrarily applicable.

Opposition is defined as the relation between two stimuli which differ from a particular point along the same dimension to the same degree (e.g., cool is opposite to warm, freezing is opposite to boiling).

The frame of comparison is involved whenever one event is responded to in terms of a nonequal and nonopposite quantitative or qualitative relation along a specific dimension (“greater than” is an example of responding according to this frame).

Responding according to relational frames involves the following three properties: mutual entailment, combinatorial entailment, and the transformation of function. We will now briefly outline each of these properties in turn.

Mutual entailment: In a given context, if a stimulus A bears a relationship to another stimulus B, then a further derived relation between B and A is mutually entailed. The type of relation entailed depends upon the type of trained relation between A and B (Hayes, 1994). For instance, if the stimulus A bears a "same as" relation to the stimulus B, then the relation "B is the same as A" is entailed. However, trained and mutually entailed relations are not always identical. For instance, given "A is more than B", a "less than" relation is mutually entailed between B and A (i.e., B is less than A).

Combinatorial entailment: In a given context, if a stimulus A bears a relation to B, and B bears a further relation to C, then a relation between A and C is derived by combinatorial entailment. The nature of this derived relation depends on the nature of the trained relations. For example, if A is more than B and B is more than C, then a "more than" relation between A and C is derived by combinatorial entailment (i.e., A is more than C) and a "less than" relation is entailed between C and A (i.e., C is less than A).

Transformation of function: If a stimulus A is related to another stimulus B, and A acquires a psychological function, then in certain contexts the stimulus functions of B will be transformed in accordance with the A-B relation. For example, if A is "more than" B, and A actualizes fear, then B will actualize less fear than A (Hayes, 1994; see also Dymond & Barnes, 1995; Roche & Barnes, 1997; Roche et al., 2000, for empirical evidence to this effect).

Measures of Derived Relational Responding

The current study is concerned in particular with response latencies to the derived relations of Same, Opposite, More Than, and Less Than. Response latency represents an additional measure of derived relational responding that has received extremely little attention within the relational frame research literature. Before we outline the current research in greater detail, therefore, it is necessary to first consider the use of response latency as a behavioral measure of derived relational responding more generally.

In studies of stimulus equivalence, performance is usually measured in terms of a percentage correct criterion. For instance, subjects may be expected to respond correctly on 9 tasks from every block of 10 (e.g., Dymond & Barnes, 1994). Typically, however, subjects will require multiple exposures to an equivalence test before satisfying this percentage correct criterion. In effect, the emergence of the equivalence effect usually occurs across testing sessions in a transition from inaccurate to accurate responding (Spencer & Chase, 1996). The percentage correct criterion has dominated as the criterion of choice for measuring equivalence performances since the outset. It has also been the favored measure in the analysis of more complex features of equivalence responding, such as differences in responding to symmetry, transitivity, and combined symmetry and transitivity trials (Bush, Sidman, & deRose, 1989), the effects of nodal distance (Fields & Verhave, 1987), and the relatedness of stimuli in equivalence classes (Fields, Adams, Verhave, & Newman, 1990).

Despite the overwhelming popularity of the response accuracy measure, some researchers have recorded other dependent measures of relational responding. Examples of these measures include the number of training trials required for particular relations to emerge (Dube, Green, & Serna, 1992; Fields et al., 1990; Kennedy, 1991) and subject estimation of reinforcer probability (see Pilgrim & Galizio, 1996). Response latency has also been used as a measure of performance in several studies of derived relational responding (e.g., Bentall, Dickins, & Fox, 1993; Steele & Hayes, 1991; Wulfert & Hayes, 1988). For example, Bentall et al. (1993) showed that subjects responded more quickly on trials for directly trained relations than on trials for derived relations. In another study, Spencer and Chase (1996) found that response speed (the inverse of response latency) was proportional to nodal distance and differed significantly across symmetry and transitivity probes but not across probes for transitivity and combined symmetry and transitivity. Indeed, Spencer and Chase (1996) argued that response latency is a more sensitive measure of relational responding than response accuracy because differences in response latency across trial types can remain when response accuracy has stabilized.

Only one study to date has examined response latencies to derived relations other than equivalence (Steele & Hayes, 1991). In that study, Steele and Hayes (1991) measured mean response latencies to the derived relations of Same and Opposite. These researchers found that response latencies for combinatorial entailment involving two trained Same relations (e.g., subjects were trained: A Same as B and B Same as C) were lower than combinatorial entailment involving two trained Opposite relations (e.g., A Opposite to B and B Opposite to C). In order to explain this finding, Steele and Hayes drew attention to the fact that Same relations do not give rise to novel relation types at higher level of complexity. For instance, two Same relations combine to entail yet another Same relation (e.g., A Same as B and B Same as C entails C Same as A). Same relations differ from Opposite relations in this respect, however. More specifically, two Opposite relations combine to entail a Same relation (i.e., A Opposite to B and B Opposite to C entails C *Same as A*). In effect, Steele & Hayes (1991) associated the combinatorial entailment of relations that are different to those trained (i.e., Opposite) with extended response latencies.

Steele and Hayes examined the foregoing effect at the level of combinatorial entailment. This was necessary because Same, Opposite, and Different do not give rise to novel relations at the level of mutual entailment (i.e., If A is the Same as, Opposite to, or Different from B, then B is the Same as, Opposite to, or Different from A, respectively). In effect, the analysis of derived relations that are different to those trained was conducted at a relatively high level of complexity in the Steele and Hayes study. In order to simplify such an analysis in terms of response latency, the current study employed the comparison relations of More Than and Less Than. Both of these relations give rise to novel relations at the level of mutual entailment (i.e., if A is More Than or Less Than B, then B is Less

Than or More Than A, respectively). Employing these relations, therefore, allows for the differentiation of various derived relations in terms of response latency while also circumventing the possible confounding effects of nodal distance encountered at the level of combinatorial entailment. Moreover, if the derivation of a relation different than that trained determines response latency to any extent, then this effect should be observable at the level of mutual entailment alone.

Experiment 1 examined differences in response latencies across probes for mutually entailed relations that were the same as (i.e., Same and Opposite relations) and different from (i.e., More Than and Less Than relations) those trained. Relational Frame Theory predicts that, as the latter is a more complex form of relational responding, higher response latencies should be observed in this case than in the former. To test this, subjects were first exposed to nonarbitrary relational pretraining for the relations of Same, Opposite, More Than, and Less Than. Subjects were then exposed to More Than/Less Than and Same/Opposite arbitrary relational training followed by testing for mutually entailed relations. Response latencies to all four probe types were then compared. Experiment 2 used the response latency measure to test the RFT idea that derived relational responding is a learned behavior that should be subject to relatively immediate practice effects.

Experiment 1

Subjects

Three unpaid female volunteers participated in Experiment 1. Subjects were between 21 and 23 years of age and were registered students at the National University of Ireland. Subjects were recruited through personal contacts. None of the subjects had any knowledge of RFT or the derived relational responding literature more generally.

Apparatus

Subjects were seated in a small experimental room (6' x 6') before an Apple Macintosh® computer (Performa 630) with a 14-inch display on which all training and testing tasks were presented. The stimuli were black and presented on the computer screen with a gray background. Presentation of the stimuli was controlled by the experiment generating software application *PsyScope* (Cohen, MacWhinney, Flatt, & Provost, 1993; see also Roche, Stewart, & Barnes-Holmes, 1999), which recorded all responses and response latencies. *PsyScope* provides a minimum of 16 ms accuracy (the temporal resolution of the Macintosh operating system) in the measurement of response latency.

Four stimuli, each consisting of a string of six characters (i.e. ??????, !!!!!, ***** , %%%%%%%%%), were used as the contextual stimuli and randomly assigned the roles of Same, Opposite, More Than, and Less Than for each subject. Ten nonsense syllables were employed (CUG, MEL, ZID, DAX, PAF, ZER, BAL, VEK, MAV, MIM) during the relational training and

testing phases. These stimuli were randomly assigned to their roles as samples and comparisons for each subject (six in the More Than/Less Than training and testing stage and four in the Same/Opposite training and testing stage) and are labeled, in the interests of clarity, using the alphanumeric; A1, B1, B2, Y1, N1, N2, C1, D1, D2, E1.

Procedure

All trials were presented on the computer monitor. On any task, the contextual cue appeared in the center top third of the screen. After a 1-s delay the sample stimulus appeared in the middle of the screen. Following a further 1-s delay, the comparison stimuli appeared in a line at the bottom of the screen. The screen position of the comparisons (i.e., left, middle, or right) was counterbalanced across trials. The contextual cue, sample, and comparisons remained on the screen together until a response was recorded. Feedback followed responses on all training trials which were followed in turn by an intertrial interval (i.e., the screen remained blank for 2.5 s). Feedback was provided during all training phases. Following a correct response the screen cleared and the word "Correct" appeared accompanied by a beep from the computer. Following an incorrect response the word "Wrong" appeared with no sound. During the testing trials, no feedback was provided (i.e., following a subject's response, the computer proceeded directly to the intertrial interval).

The Z, V, and M keys on the computer keyboard were colored red and designated as response keys. At the beginning of each training and testing phase, the following instructions were displayed on the computer screen:

In a moment some images will appear on this screen. Your task is to first look at the image at the top of the screen, then look at the image in the middle of the screen and on the basis of these two objects choose one of the images at the bottom of the screen by pressing one of the red buttons in front of you. So, if you want to choose an image on the left, press the red button on the left. If you want to choose an image in the middle, press the red button in the middle. And if you want to choose an image on the right then you should press the red button on the right. Hit any key when you are ready to begin.

Same/Opposite pretraining. The purpose of relational pretraining is to establish contextual control functions (e.g., Same and Opposite) for a series of arbitrary stimuli to be used during the subsequent arbitrary relational training stage. The current relational pretraining phase closely followed that reported by Steele and Hayes (1991). During each task in this phase, the contextual cue for Same or Opposite appeared in the center top third of the screen, the sample stimulus appeared in the middle of the screen, and three comparison stimuli appeared in a line at the bottom of the screen. On all tasks the sample and comparison stimuli were related to each other along some physical dimension. For example, one set of comparisons for this stage consisted of a long line, a medium line, and a short line. Thus, given a short-line sample, in the presence of the Opposite contextual cue (e.g.,

??????), responses to the long-line comparison were reinforced. However, given the Same contextual cue (e.g., !!!!!) and a short line, responses to the short-line comparison were reinforced. In this way contextual control was established. Four tasks constituted one problem set (e.g., Same/long line-long line, Same/short line-short line, Opposite/long line-short line, Opposite/short line-long line).

During this phase, the subjects were exposed to three problem sets (Sets 1-3) in a quasi-random order. Each of the four tasks from the first problem set were presented twice in a quasi-random order across a block of eight trials. This procedure was then repeated for the remaining two problem sets. Following these 24 trials, 2 trials from each of the three problem sets were presented in a quasi-random order (i.e., 6 trials). In effect, there were 30 training trials in all (i.e., each problem set consisted of 10 trials) and subjects were required to produce 9 out of 10 correct responses for each of the three problem sets. If subjects failed to satisfy the mastery criterion, they were reexposed to the training and testing cycle up to a maximum of three times. All 3 subjects satisfied the mastery criterion within four exposures to the pretraining phase before proceeding to the testing stage.

The testing procedure was identical to the training procedure with the difference that feedback was not provided. The testing phase consisted of three novel problem sets (Sets 4-6) presented in a quasi-random order. Each of the four tasks from the first problem set were presented twice in a quasi-random order across a block of eight trials. This procedure was then repeated for the remaining two problem sets. Following these 24 trials, 2 trials from each of the three problem sets were presented in a quasi-random order (i.e., 6 trials). In effect, there were 30 testing trials in all (i.e., each problem set consisted of 10 trials) and subjects were required to produce 9 out of 10 correct responses for each of the three problem sets.

If subjects failed to satisfy the mastery criterion on the first exposure to the testing phase, they were retrained on Problem Sets 1-4 in a quasi-random order. Each of the four tasks from the first problem set were presented twice in a quasi-random order across a block of eight trials. This procedure was then repeated for the remaining two problem sets. Following these 32 trials, 2 trials from each of the three problem sets were presented in a quasi-random order (i.e., 8 trials). Thus, there were 40 training trials in all (i.e., each problem set consisted of 10 trials). Subjects were required to produce 9 out of 10 correct responses for each of the four problem sets. If subjects failed to satisfy the mastery criterion they were reexposed to the training and testing cycle up to a maximum of three times. If subjects satisfied the mastery criterion they proceeded to the second testing phase.

During the second testing phase subjects were retested on Problem Sets 5 and 6 plus a completely novel Set 7. Each of the four tasks from the first problem set were presented twice in a quasi-random order across a block of eight trials. This procedure was then repeated for the remaining two problem sets. Following these 24 trials, 2 trials from each of the three

problem sets were presented in a quasi-random order (i.e., 6 trials). Thus, there were 30 testing trials in all (i.e., each problem set consisted of 10 trials) and subjects were required to produce 9 out of 10 correct responses for each of the three problem sets. No subject failed at this level of testing.

More Than/Less Than pretraining. All 3 subjects achieved the mastery criterion on the Same/Opposite testing phase before proceeding to More Than/Less Than pretraining. The procedure for the More Than/Less Than relational pretraining stage was similar to that used by Dymond and Barnes (1995). During More Than/Less Than pretraining, contextual functions were established for two contextual stimuli. On each task, the contextual cue for More Than or Less Than appeared in the center top third of the screen, the sample stimulus appeared in the middle of the screen, and two comparison stimuli appeared in a line at the bottom of the screen. On all tasks sample and comparison stimuli were related to each other along some physical dimension. For example, one set of comparisons for this stage consisted of a four-circle comparison and a two-circle comparison. Thus, given a three-circle stimulus as a sample, in the presence of the More Than contextual cue (e.g., %%%%), responses to the four-circle comparison were reinforced. However, in the presence of the Less Than cue (e.g., *****), responses to the two-circle comparison were reinforced. In this way contextual control was established for two arbitrary stimuli. Two tasks constituted one problem set (e.g., More Than/three circles-four circles, Less Than/three circles-two circles).

During this phase, subjects were exposed to three problem sets (Sets 1-3) in a quasi-random order. Each of the four tasks from the first problem set were presented four times in a quasi-random order across a block of eight trials. This procedure was then repeated for the remaining two problem sets. Following these 24 trials, 2 trials from each of the three problem sets were presented in a quasi-random order (i.e., 6 trials). Thus, there were 30 training trials in all (i.e., each problem set consisted of 10 trials) and subjects were required to produce 9 out of 10 correct responses for each of the three problem sets. If subjects failed to satisfy the mastery criterion, they were reexposed to the training and testing cycle up to a maximum of three times. All three subjects satisfied the mastery criterion within four exposures to the pretraining phase before proceeding to the testing phase.

The testing procedure was identical to the training procedure with the difference that feedback was not provided. The testing phase consisted of three novel problem sets (Sets 4-6) presented in a quasi-random order. Both tasks from the first problem set were presented four times in a quasi-random order across a block of eight trials. This procedure was then repeated for the remaining two problem sets. Following these 24 trials, both trials from each of the three problem sets were presented in a quasi-random order (i.e., 6 trials). In effect, there were 30 testing trials in all (i.e., each problem set consisted of 10 trials) and subjects were required to produce 9 out of 10 correct responses for each of the three problem sets.

If subjects failed to satisfy the mastery criterion on the first exposure

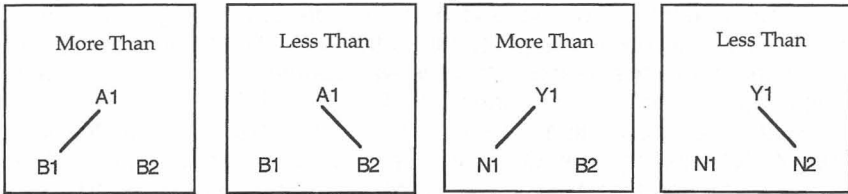
to the testing phase, they were retrained on Problem Sets 1-4 in a quasi-random order. Both tasks from the first problem set were presented four times in a quasi-random order across a block of eight trials. This procedure was then repeated for the remaining two problem sets. Following these 32 trials, 2 trials from each of the three problem sets were presented in a quasi-random order (i.e., 8 trials). Thus, there were 40 training trials in all (i.e., each problem set consisted of 10 trials). Subjects were required to produce 9 out of 10 correct responses for each of the four problem sets. If subjects failed to satisfy the mastery criterion they were reexposed to the training and testing cycle up to a maximum of three times. If subjects satisfied the mastery criterion they proceeded to the second testing phase.

During the second testing phase subjects were retested on Problem Sets 5 and 6 plus a completely novel Set 7. Both tasks from the first problem set were presented four times in a quasi-random order across a block of eight trials. This procedure was then repeated for the remaining two problem sets. Following these 24 trials, both trials from each of the three problem sets were presented in a quasi-random order (i.e., six trials). Thus, there were 30 testing trials in all (i.e., each problem set consisted of 10 trials) and subjects were required to produce 9 out of 10 correct responses for each of the three problem sets. No subject required further training or testing beyond this level.

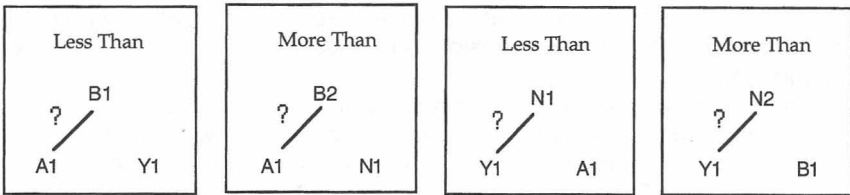
More Than/Less Than arbitrary relational training and testing. Following relational pretraining subjects were exposed to four arbitrary relational training tasks presented in a quasi-random order within blocks of 40 trials (i.e., 10 exposures to each task within a 40-trial block with no more than two successive exposures to any task). A diagrammatic representation of all training and testing tasks can be seen in Figure 1. The stimuli were six randomly chosen nonsense syllables (i.e., A1, B1, B2, Y1, N1, N2) and two contextual stimuli, which had acquired the contextual functions of More Than and Less Than during pretraining. The four training tasks were: More Than/A1-B1, Less Than/A1-B2, More Than/Y1-N1, and Less Than/Y1-N2 (see Figure 1). Subjects were required to produce nine correct responses across the 10 exposures to each of the four tasks within a 40-trial block. If subjects failed to satisfy the mastery criterion, they were reexposed to the training phase up to a maximum of three times. If subjects satisfied the mastery criterion they were exposed immediately to More Than/Less Than arbitrary relational testing.

The testing phase probed for the following mutually entailed relations: More Than/B2-A1, Less Than/B1-A1, Less Than/N1-Y1, More Than/N2-Y1 (see Figure 1). Tasks were presented without feedback 10 times each across a block of 40 trials in a quasi-random order (i.e., no more than two successive exposures to any task). If subjects failed to produce nine correct responses across the 10 exposures to each of the four tasks, they were reexposed to the relational training and testing cycle up to a maximum of three times. All subjects satisfied the mastery criterion within three exposures to the test phase.

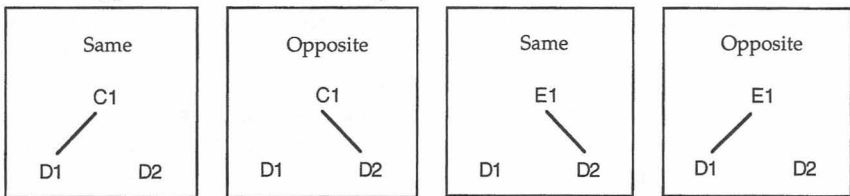
More Than/Less Than Relational Training



More Than/Less Than Relational Testing



Same/Opposite Relational Training



Same/Opposite Relational Testing

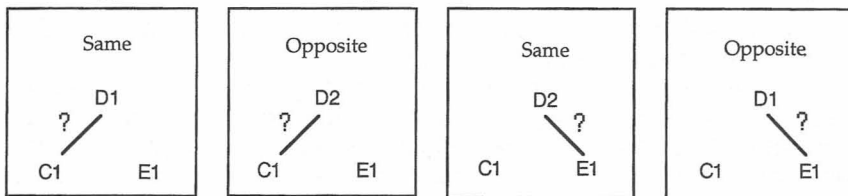


Figure 1. A diagrammatic representation of all arbitrary relational training and testing tasks used in Experiments 1 and 2. Solid lines indicate trained relations. Solid lines accompanied by a question mark indicate predicted derived relations during test phases.

It is important to note that the test tasks were carefully designed to eliminate extraneous sources of control during testing. First, the N1 stimulus, which was chosen as a correct comparison on the More Than/Y1-N1 training task, was also presented as an *incorrect* comparison on the More Than/B2-A1 probe (see Figure 1). Presenting N1 as the incorrect comparison on this latter task ensured that subjects did not

respond directly to the contextual cue as a sample (i.e., always choose N1 in the presence of the More Than cue). More specifically, because N1 was a correct choice in the presence of the More Than cue on a training task, attention to the sample stimulus was required in order to correctly choose the A1 comparison during the More Than/B2-A1 probe.

Second, B2 was included as an unreinforced comparison on the More Than/Y1-N1 training task. Creating this overlap across the A-B and Y-N training tasks precluded the possibility that subjects could respond during test probes simply by matching samples and comparisons that had previously appeared together in a trigram configuration. During the More Than/B2-A1 test probe, therefore, both the correct A1 comparison and the incorrect N1 comparison had previously appeared together with B2 in training tasks (B2 was the incorrect comparison in the More Than/Y1-N1 training task).

Third, in an earlier version of the current study, 1 subject noticed and reported that sample stimuli from the training phase always appeared as correct comparisons during testing. In the current study, however, the two test tasks Less Than/B1-A1 and Less Than/N1-Y1 controlled for this possibility. More specifically, during the Less Than/B1-A1 task, Y1 was presented as the incorrect comparison. In effect, *both* the A1 and Y1 comparisons presented on this task had been presented as samples on previous training tasks (i.e., More Than/A1-B1 and More Than/Y1-N1, respectively). Similarly, during the Less Than/N1-Y1 task, A1 was presented as the incorrect comparison. Here again, both A1 and Y1 comparisons had been presented as samples on previous training tasks, thereby precluding the possibility that subjects could respond reliably by choosing only those comparisons that had served as samples during the training phase.

Having successfully completed More Than/Less Than arbitrary relational training and testing, Subjects 1 and 2 were exposed to Same/Opposite arbitrary relational training and testing. Subject 3, however, was required to successfully complete Same/Opposite arbitrary relational training and testing before proceeding to More Than/Less Than arbitrary relational training and testing.

Same/Opposite arbitrary relational training and testing. During this phase, subjects were exposed to four training tasks: Same/C1-D1, Opposite/C1-D2, Same/E1-D2, Opposite/E1-D1 (see Figure 1). These tasks were presented 10 times each in a block of 40 trials in a quasi-random order (i.e., no more than two successive exposures to any task). Subjects were required to choose the correct comparison 9 times out of every 10 exposures to a task before proceeding to Same/Opposite arbitrary relational testing. If subjects failed to satisfy the mastery criterion they were reexposed to the training phase up to a maximum of three times. If subjects satisfied the mastery criterion they were exposed immediately to Same/Opposite arbitrary relational testing.

During testing, subjects were presented with probes for the following derived stimulus relations; Same/D1-C1, Opposite/D2-C1, Same/D2-E1, Opposite/D1-E1 (see Figure 1). Probes were presented 10 times each in

a quasi-random order across a block of 40 trials (i.e., no more than two successive exposures to any task). If subjects failed to produce nine correct responses across the 10 exposures to each of the four tasks, they were reexposed to the relational training and testing cycle up to a maximum of three times. All three subjects satisfied the mastery criterion on the first exposure to the testing phase.

Results

Subject 1 required three exposures to the More Than/Less Than test, achieving 22.5%, 32.5%, and 92.5% accuracy respectively, on successive exposures. Subject 2 also required three exposures to the More Than/Less Than test achieving 25%, 0%, and 100% accuracy, on successive exposures. Subject 3 demonstrated the required performance on the More Than/Less Than test on her first exposure (97.5% accuracy). Subjects 1, 2, and 3 achieved the mastery criterion on their first exposure to the Same/Opposite test (100%, 97.5%, and 100% accuracy, respectively). Figure 2 shows response latencies in milliseconds to each successive probe, and the median for each block of 10 probes (i.e., each quartile) during the Same/Opposite and More Than/Less Than tests. Medians are more representative of the current data set than means because variance within the current data set was relatively large. Median response latencies to the More Than/Less Than probes were greater than the median response latencies to Same/Opposite probes for all 3 subjects.

For all 3 subjects, the median response latency in the first quartile of Same/Opposite probes was greater than the median response latency in the fourth quartile. For Subjects 1 and 3, the median response latency in the first quartile of More Than/Less Than probes was greater than the median response latency in the fourth quartile. Relative differences in response latencies between the four quartiles were more varied.

For the purpose of statistical analysis, all response latency data were log transformed. Log transformations serve to reduce skew and kurtosis in data distributions with large variance, thereby rendering subsequent statistical analyses more conservative. A two-way analysis of variance was conducted to examine (a) the differences in response latencies during the Same/Opposite and More Than/Less Than tests, (b) the decrease in response latencies across quartiles, and (c) any interaction between these two effects (see Table 1). The observed differences in response latencies across the Same/Opposite and More Than/Less Than tests were significant at $p \leq 0.05$ for all 3 subjects. The decreases in response latencies across quartiles were also significant ($p \leq 0.01$) for Subjects 1 and 3. For all 3 subjects, there was no significant interaction between relation type (i.e., More Than/Less Than and Same/Opposite) and quartiles and therefore no separate statistical analysis of the relative differences between quartiles within the stimulus sets was necessary. In effect, for Subjects 1 and 3, response latencies to the Same/Opposite and More Than/Less Than probes tracked each other as they decreased

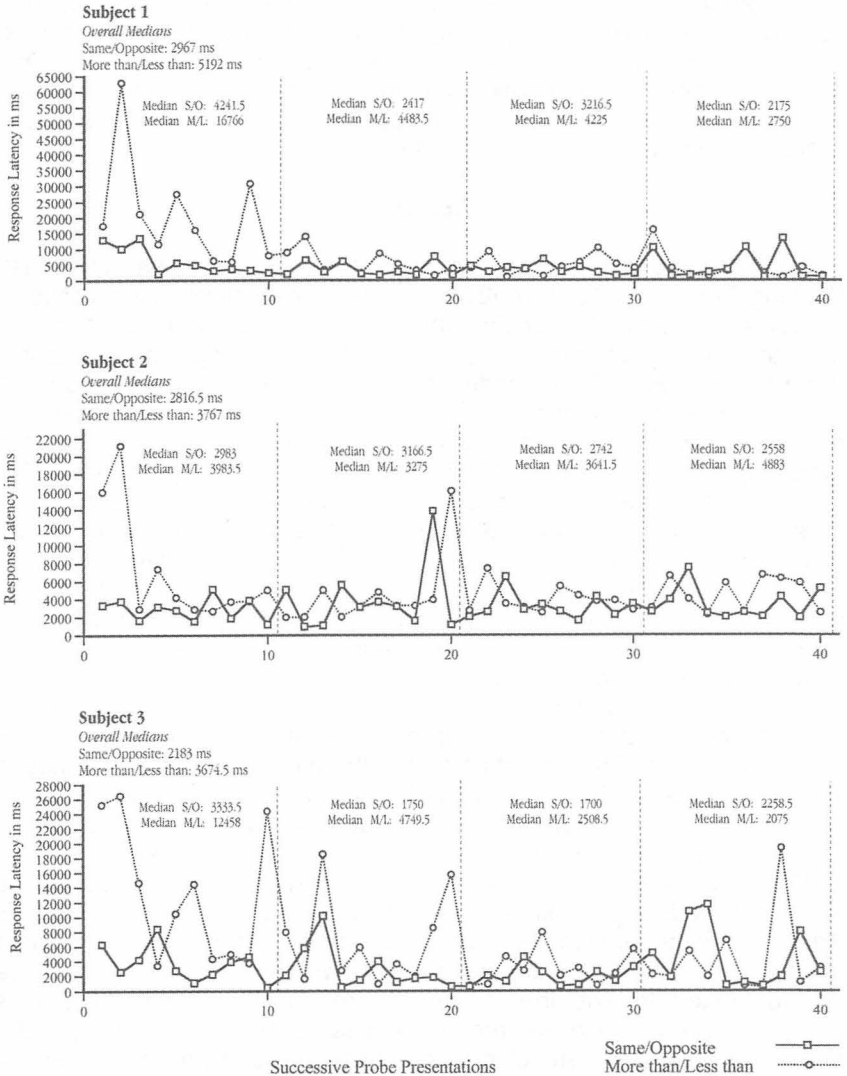


Figure 2. A graphical representation of response latencies to probes for mutually entailed relations of More Than, Less Than, Same, and Opposite in Experiment 1. Response latencies are represented in milliseconds along the Y axis and successive probes are represented on the X axis in the original order of presentation. Vertical hashed lines represent divisions between successive quartiles. Median response latencies to Same/Opposite and More Than/Less Than probes are provided for each quartile. The acronyms S/O and M/L refer to the Same/Opposite and More Than/Less Than tests, respectively. Overall median response latencies to Same/Opposite and More Than/Less Than test probes are also provided.

Table 1

ANOVA Values on Log Transformed Response Latencies for Relation Type, Quartile, and Interaction Between These Two Variables in Experiment 1

		<i>df</i>	<i>F</i>	<i>p</i>
Subject 1	Relation type	1	8.299	0.052
	Quartile	3	8.767	0.0001
	AB	3	2.315	0.083
Subject 2	Relation type	1	9.178	0.0034
	Quartile	3	0.292	0.8308
	AB	3	0.783	0.5073
Subject 3	Relation type	1	8.025	0.006
	Quartile	3	4.311	0.0075
	AB	3	2.129	0.104

across test sessions. For Subject 2, the response latency differential to probes for Same/Opposite and More Than/Less Than remained constant from the beginning to the end of the test sessions.

Experiment 2

Experiment 1 found that response latencies to probes for More Than/Less Than relations were significantly greater than latencies to probes for Same/Opposite relations. The findings of Experiment 1 also provide support for the operant nature of the RFT account insofar as response latencies decreased across quartiles from the beginning to the end of the test sessions for all 3 subjects (this decrease was significant for Subjects 1 and 3). A more important question from the RFT perspective, however, is whether or not this decrease in response latencies would have *generalized* to novel stimulus sets. More specifically, if the RFT view is correct, we might expect that providing subjects with an exemplar of the More Than, Less Than, Same, and Opposite relations, should increase proficiency in responding to probes for these relations using novel stimulus sets. In effect, an extension of subjects' histories of reinforcement for deriving each of the four relations should lead to decreased response latencies on probes for these relations when novel stimulus sets are used. Experiment 2 was designed to examine this possibility.

In Experiment 2, 4 subjects were exposed to Same/Opposite and More Than/Less Than pretraining, as in Experiment 1. Subjects 4 and 5 were then exposed to the Same/Opposite arbitrary relational training and testing followed by Same/Opposite arbitrary relational training and testing with a novel set of stimuli. Subjects 6 and 7 were exposed to the More Than/Less Than training and testing followed by More Than/Less Than arbitrary relational training and testing using a novel stimulus set.

Subjects

One male and three female unpaid volunteers between 18 and 21

years of age were recruited through personal contacts. Subjects were registered students at the National University of Ireland. None of the subjects had knowledge of Relational Frame Theory or the derived relational responding literature more generally.

Apparatus

The materials were identical to those used in Experiment 1 except that a further 10 nonsense syllables were used in this experiment as a novel stimulus set during the relational training and testing phases (i.e., DOF, CET, VEP, LEB, ARL, VOL, YAT, CUX, BEH, SEP). These stimuli were randomly assigned to their roles as samples and comparisons for each subject (six in the More Than/Less Than training and testing stage and four in the Same/Opposite training and testing stage).

Procedure

The procedure for training and testing relational responding in Experiment 2 was the same as in Experiment 1. Only the order of presentation of the training and testing phases differed across the two experiments. All subjects were first exposed to the Same/Opposite and More Than/Less Than relational pretraining stages. Following this, Subjects 4 and 5 were immediately exposed to the Same/Opposite arbitrary relational training followed by testing for the mutually entailed Same/Opposite relations. These subjects were then exposed to Same/Opposite arbitrary relational training and testing using a novel stimulus set (see Figure 1 for all training and testing tasks). Subjects 6 and 7 were exposed to relational pretraining for Same, Opposite, More Than, and Less Than followed by More Than/Less Than arbitrary relational training and testing. These subjects were then exposed to the More Than/Less Than arbitrary relational training and testing using the novel stimulus set.

Results

Subjects 4 and 5 achieved the mastery criterion on their first exposure to the Same/Opposite test (97.5% and 100% accuracy, respectively). Subject 6 required two exposures to the More Than/Less Than test, achieving 52.5% and 100% accuracy on successive exposures. Subject 7 demonstrated the required performance on the More Than/Less Than test on her first exposure (100% accuracy). All subjects achieved the mastery criterion on their first exposure to the Same/Opposite (Subjects 4 and 5) or More Than/Less Than (Subjects 6 and 7) tests using the novel stimulus sets (100% accuracy in all cases).

Figure 3 shows the response latencies to successive probes during the tests using the original and the novel stimulus sets. For all 4 subjects, median response latencies to the test using the original stimulus set were greater than the median response latencies to the test using the novel stimulus set.

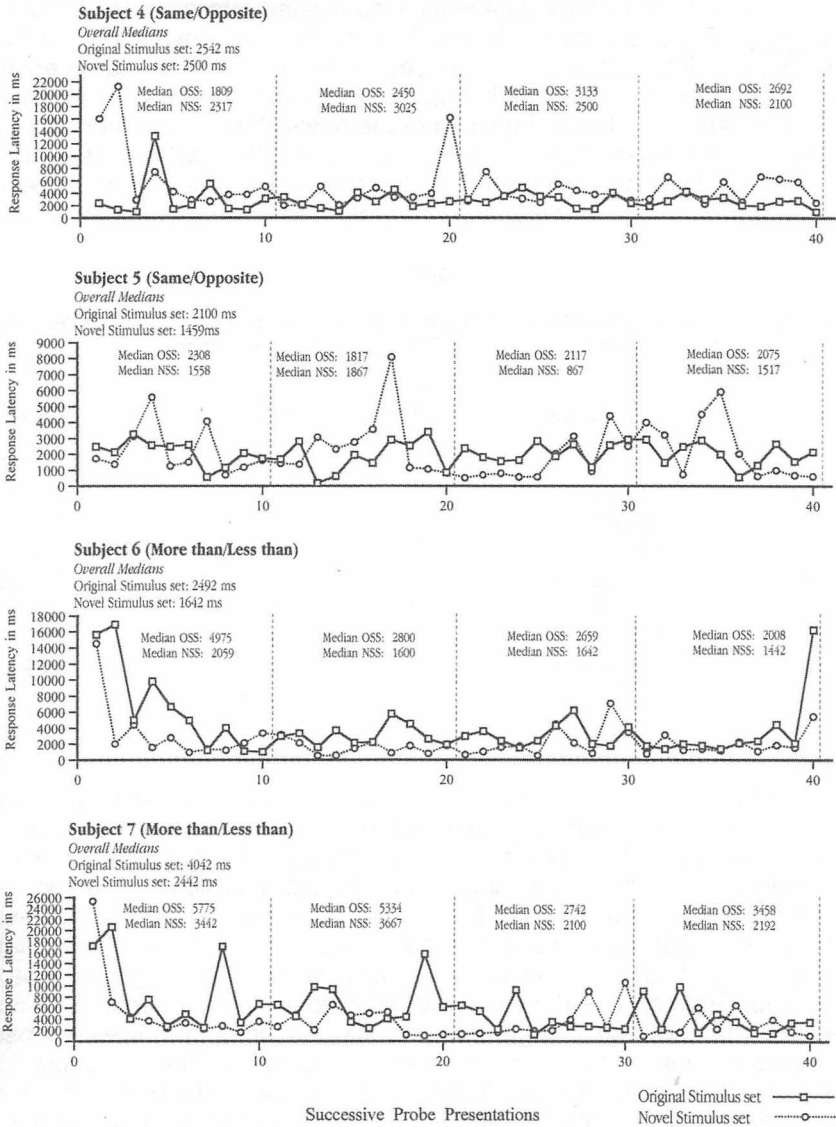


Figure 3. A graphical representation of response latencies to probes during the first successful test performance using the original stimulus set and the novel stimulus set in Experiment 2. Response latencies are represented in milliseconds along the Y axis and successive probes are represented on the X axis in the original order of presentation. Vertical hashed lines represent divisions between successive quartiles. Median response latencies to probes during the test using the original stimulus set and the novel stimulus set are provided for each quartile. The acronyms OSS and NSS refer to the original stimulus set and the novel stimulus set, respectively. Overall median response latencies for both test performances are also provided.

The response latency data were also divided into quartiles for both stimulus sets. For Subjects 5, 6, and 7 the median response latency in the first quartile of the original stimulus set probes was greater than median response latency in the fourth quartile. For all 4 subjects the median response latency in the first quartile of the novel stimulus set probes was greater than median response latency in the fourth quartile (see Figure 3). In effect, response latencies decreased from the beginning to the end of test sessions, with only one exception (Subject 4, original stimulus set).

Table 2

ANOVA Values on Log Transformed Response Latencies for Stimulus Set, Quartile, and Interaction Between These Two Variables in Experiment 2

		<i>df</i>	<i>F</i>	<i>p</i>
Subject 4	Stimulus set	1	0.154	0.6962
	Quartile	3	0.323	0.8086
	AB	3	0.314	0.8151
Subject 5	Stimulus set	1	0.297	0.5872
	Quartile	3	0.224	0.8793
	AB	3	1.41	0.2468
Subject 6	Stimulus set	1	12.879	0.0006
	Quartile	3	2.124	0.1047
	AB	3	0.266	0.8497
Subject 7	Stimulus set	1	7.776	0.0068
	Quartile	3	3.122	0.0311
	AB	3	0.577	0.632

For the purpose of statistical analysis, all response latency data was log transformed. A two-way analysis of variance was conducted to examine (a) the differences in response latencies during the tests using the original and novel stimulus sets, (b) the decrease in response latencies across quartiles, and (c) any interaction between these two effects (see Table 2). The observed difference in response latencies across the tests using the original and novel stimulus was significant at $p \leq 0.01$ for Subjects 6 and 7 (i.e., who received the More Than/Less Than training and testing). The decreases in response latencies across quartiles were significant ($p \leq 0.05$) only for Subject 7. For all 4 subjects, there was no significant interaction between stimulus set (i.e., original and novel) and quartile. Thus, no separate statistical analysis of the relative differences between quartiles within the stimulus sets was necessary.

General Discussion

The current study extends upon the earlier finding of Steele and Hayes (1991) that response latencies increase when responding correctly to probes involves more than one type of relation (e.g., responding to both Same and Opposite relations). Experiment 1 of the current study built

upon this finding by comparing response latencies to probes for mutually entailed relations that were the same as those trained (e.g., A same as B mutually entails B same as A) with response latencies to probes for relations that produced novel relations at the level of mutual entailment (A more than B mutually entails B *less than* A). It was found that response latencies to probes for mutually entailed relations were greater when these relations were different from those trained. More specifically, response latencies to More Than/Less Than probes were significantly greater than response latencies to probes for Same/Opposite relations. The current findings, therefore, appear to parallel those of Steele and Hayes insofar as the more complex probe tasks (i.e., More Than/Less Than) yielded greater response latencies.

Experiment 2 examined the possibility that the greater response latencies to the More Than/Less Than probes were caused, at least in part, by differences in the extended reinforcement histories for responding in accordance with exemplars of the relations of Same, Opposite, More, and Less. It was found that response latencies to novel More Than/Less Than probes could be decreased significantly with training on an exemplar of these relations. The findings of Experiments 1 and 2, therefore, support the RFT predictions that derived relational responding performances should (a) improve with extended training and (b) generalize to novel sets of stimuli.

Interestingly, in Experiment 2 response latencies to probes for More Than/Less Than relations decreased significantly from the original to the novel stimulus set (i.e., Subjects 6 and 7), whereas response latencies to probes for Same/Opposite relations did not (i.e., Subjects 4 and 5). This outcome may seem to run counter to the RFT view that derived relational performances should improve with exposure to multiple exemplars. However, as outlined earlier, Same and Opposite relations are not complex at the level of mutual entailment insofar as relations derived at this level are always the same as those trained. In effect, Same and Opposite relations are functionally identical at the level of mutual entailment. It is consistent with RFT, therefore, to expect that performances on Same/Opposite tasks will be at an optimal level for most verbally-able adults as they enter the laboratory. One feature of this optimal performance should be that response latencies on any test for Same/Opposite responding already approach their minimum. Thus, the absence of a significant decrease in response latencies to Same/Opposite probes, from the original to the novel stimulus sets in Experiment 2, was likely due to the already low response latencies to these probe types within the original stimulus set (i.e., a floor effect). Indeed, Figures 2 and 3 illustrate that response latencies were low and stable on the *first* exposure to Same/Opposite probes for all subjects in both Experiments 1 and 2.

The reader may attempt to explain subjects' test performances in both Experiments 1 and 2 in terms of the unusual configuration of tasks used during the current relational training and testing phases. More specifically,

it could be argued that subjects responded to all test probes simply by matching samples and comparisons that “went together” during training (i.e., by ignoring the contextual cue and responding on the basis of symmetry relations between samples and comparisons). For instance, during the More Than/B2-A1 probe, subjects may have paired the B2 sample and A1 comparison because they “went together” during the Less Than/A1-B2 training task. A similar mode of responding was also possible during the remaining More Than/Less Than and Same/Opposite probes.

The idea that subjects ignored the contextual cues during testing raises two problems. First, it is not clear why performances should come under clear contextual control during relational pretraining and testing as well as during arbitrary relational training, but not during the arbitrary relational test probes. Indeed, during post-hoc interviews all of the subjects referred to the contextual cues using appropriate English words, such as “More,” “Greater,” “Less,” “Same,” and “Opposite,” and clearly indicated that they had used these cues in making their comparison choices.

Second, if subjects did indeed ignore the contextual cues during testing, this would render the More Than/Less Than and Same/Opposite tests functionally identical. In effect, we would not expect to see any reaction time differences across the two test performances, as we did in the current study. Thus, it served as a convenient control that subjects could respond during *all* probes across both the More Than/Less Than and Same/Opposite tests on the basis of symmetrical relations between samples and comparisons.

The foregoing highlights an inherent difficulty in designing training task configurations for the analysis of multiple stimulus relations. The very nature of certain derived relations constrains the configurations of tasks that can be used to train them. Consider, for instance, the two training tasks More Than/A1-B1 and Less Than/A1-B2. Together these two relations allow subjects to derive that both B1 and A1 are More Than B2. Thus, during the More Than/B2-A1 probe we could not present B1 as the incorrect comparison because B1 was also a correct choice on this task. Instead, N1 was used as the incorrect comparison, thereby constraining the control that we could exert over task configurations across the More Than/Less Than and Same/Opposite training phases (see procedure section for why this use of N1 also protected against direct control by contextual cues over comparison choice).

Despite the various constraints on training and testing task configurations, it should be remembered that a derived relation is partly *defined* by the unique configurations of tasks required to train it. Thus, differences in task configurations across the More Than/Less Than and Same/Opposite training phases do not pose a problem conceptually. Indeed, if we could eliminate all such differences between the More Than/Less Than and Same/Opposite training tasks, the resulting derived relations would be functionally identical. Consequently, we would not expect to find any differences in performances or reaction times to probes across different tests, as we did in the current study. Of course, further

research would do well to examine the various and multiple features of relations that control performances and reaction times. Nevertheless, Experiment 1 of the current study served at least as a *prima facie* analysis of the reaction times that characterize the derived relations of Same, Opposite, More Than, and Less Than.

It is worth considering at this point one possible feature of the More Than/Less Than relations that may have impacted on response latencies during the More Than/Less Than probes. Specifically, during More Than/Less Than training trials it may have been possible for subjects to respond covertly to combinatorially derived relations between the correct and incorrect comparisons. For instance on the More Than/Y1-N1 training task subjects may have responded to the incorrect B2 comparison as Less Than (or equal to) Y1, although this response was never reinforced. In effect, subjects' responses to the correct comparisons may have come under a form of S- control, whereby the relation between the sample and the incorrect comparison could always be inferred within each task (i.e., if N1 is more than Y1 then B2 must be less than or equal to Y1). Thus, B2 may have participated in a specified relation with the Y1 sample. Consequently, the directly trained More Than/Y1-N1 relation and the inferred Less Than (or equal to)/Y1-B2 relation may have combined such that subjects derived the relations More Than/N1-B2 and Less Than/B2-N1, again without reinforcement. If a subject continued to respond to such combinatorial relations during the test phase, ambiguity might have emerged for one of the four test probes. More specifically, during the More Than/B2-A1 task, the N1 comparison may have appeared as a correct choice to a subject on the basis of the combinatorially entailed relations More Than/N1-B2 and Less Than/B2-N1. In effect, the presence of two potentially correct comparisons on this task may have impacted upon subjects' reaction times in making a choice.

Although the foregoing ambiguity may have incremented response latencies to the More Than/B2-A1 probe, it is important to remember that the ambiguity described above could only have occurred during the More Than/B2-A1 task. In contrast, subjects could respond correctly to all of the More Than and Less Than test probes at the level of mutual entailment. It would be unexpected, therefore, that subjects' responses on the More Than/B2-A1 task would come under a complex form of relational control, while responses on other tasks would be under the control of a simpler and different relational context (i.e., mutual entailment).

There are two further reasons why combinatorially entailed relations were unlikely to have controlled responses in the current studies. Firstly, had S- control been in effect during test probes for mutually entailed More Than/Less Than relations, we would not expect to have seen such consistent responding during the test probes across subjects and in accordance with the RFT predictions. Secondly, the nature of the trained More Than/Less Than relational network was such that it precluded consistent S- control during training. More specifically, the four relational training tasks did not reduce to two conditional discriminations, as might

be typical when using the matching-to-sample training and testing format. As explained in the Procedure section for Experiment 1, the training tasks did not consist of conditional discriminations because several control features were added to ensure tight relational control over subjects' responses. In effect, it was not possible for subjects to pass the training phase by producing responses under S- control alone. It would be highly surprising, therefore, if this form of control spontaneously emerged during the testing phase. In any case, even if this, or some other form of extraneous stimulus control, were to have impacted upon response latencies to the More Than/Less Than probes, it is not clear why such factors would have had a diminishing impact on response latencies across successive probes (Experiments 1 and 2) and stimulus sets (Experiment 2).

Another issue that should be considered is that response latency as a measure of derived relational responding is subject to interpretation in many different terms. Response latency may be considered a useful concurrent measure of derived relational responding performances, or even a preferred measure for differentiating performances when response accuracy has stabilized (see Spencer & Chase, 1996). The current study employed response latency as one of several possible concurrent measures of derived relational responding. It is important to understand, however, that the use of the response latency measure in no way implies an interest in response latency as an explanatory mechanism or evidence of mediating processes (i.e., as an independent variable). Moreover, this measure was not used here to cast light on problem solving strategies used by subjects during tests for derived relational responding (see Holth & Arntzen, 2000). The current study was concerned primarily with the reaction time characteristics of derived responses to relations of Same, Opposite, More Than, and Less Than, and the effects on reaction times of *histories* of exemplars in those relations.

A broader theoretical issue that is raised by these data concerns Sidman's (2000) recent assertion that the theater of controversy in the area of derived stimulus relations should be shifted from the usual philosophical and logical differences among the major theories to more empirical grounds (see also Clayton & Hayes, 1999). This is a rather perplexing statement in light of the current data and similar empirical studies, published over the last decade, that have demonstrated control in accordance with multiple stimulus relations (Barnes, Hegarty, & Smeets, 1997; Barnes & Keenan, 1993; Barnes-Holmes, Healy, & Hayes, 2000; Dymond & Barnes, 1995, 1996; Roche & Barnes, 1996, 1997; Roche et al., 2000; Steele & Hayes, 1991; Stewart, Barnes-Holmes, Roche, & Smeets, 2001). As argued by Barnes and Roche (1996) and Hayes and Barnes (1997), it is difficult to account for or even describe multiple stimulus relational control in terms of equivalence classes alone. Sidman's (1994) only attempt to address multiple stimulus relations, in terms of contextually controlled equivalence classes, was found to be wholly inadequate in that it undermined the very concept of the equivalence class itself (Barnes & Roche, 1996; Hayes & Barnes 1997).

Furthermore, Sidman's recent appeal to reinforcement contingencies as an explanation for equivalence class formation still fails to explain the data on multiple stimulus relations including those reported in the current article. For example, why would different response latencies emerge across probes for different types of relations if, as Sidman suggests, emergent responses simply "drop out" of the reinforcement contingencies (i.e., why should some mutually entailed relational responses drop out more or less rapidly than other mutually entailed responses?). To address this and many related questions requires that serious attention must be paid to the relevant histories of reinforcement for responding in accordance with nonequivalence relations, and this is precisely what RFT attempts to do. The undeniable utility of RFT in dealing with multiple stimulus relations may well explain why Sidman (2000) failed to address or even cite any of the empirical studies on multiple stimulus relations in his most recent article. We invite him once again, therefore, to join us in the empirical arena by offering an account of the current data that is clearly predicted from within his theoretical framework.

Summary and Conclusion

The findings of Experiment 1 of the current study parallel those of Steele and Hayes (1991) insofar as response latencies were found to increase with task complexity. The current findings also provide support for the RFT view that the lower response latencies observed for the Same/Opposite relations were due, at least in part, to an extended preexperimental history of reinforcement for deriving these relations. More specifically, Experiment 2 demonstrated that proficiency in arbitrarily applicable relational responding in accordance with the relations of More Than and Less Than can be enhanced by exposing subjects to multiple exemplars of these relations. It is important to remember, however, that this effect was achieved using a relatively simple laboratory intervention. Future research in the applied domain will be required to more clearly establish the power of these effects in supplementing the derived relational responding repertoires of verbally-able adults and, ultimately, establishing such repertoires where they are deficient or even entirely absent.

References

- BARNES, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record*, 44, 91-124.
- BARNES, D., HEGARTY, N., & SMEETS, P. M. (1997). Relating equivalence relations to equivalence relations: A relational framing model of complex human functioning. *The Analysis of Verbal Behavior*, 14, 1-27.
- BARNES, D., & HOLMES, Y. (1991). Radical behaviorism, stimulus equivalence, and human cognition. *The Psychological Record*, 41, 19-31.

- BARNES, D., & KEENAN, M. (1993). A transfer of functions through derived arbitrary and non-arbitrary stimulus relations. *Journal of the Experimental Analysis of Behavior*, 59, 61-81.
- 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.
- BARNES-HOLMES, D., HEALY, O., & HAYES, S. C. (2000). Relational frame theory and the relational evaluation procedure: Approaching human language as derived relational responding. In J. C. Leslie & D. E. Blackman (Eds.), *Experimental and applied analyses of human behavior* (pp. 149-180). Reno, NV: Context Press.
- BARNES-HOLMES, Y., BARNES-HOLMES, D., & CULLINAN, V. (2001). Education. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 181-196). New York: Plenum Press.
- BENTALL, R. P., DICKINS, D. W., & FOX, S. R. A. (1993). Naming and equivalence: Response latencies for emergent relations. *The Quarterly Journal of Experimental Psychology*, 46B, 187-214.
- BUSH, K. M., SIDMAN, M., & DEROSE, T. (1989). Contextual control of emergent equivalence relations. *Journal of the Experimental Analysis of Behavior*, 51, 29-45.
- COHEN, J. D., MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments and Computers*, 25(2), 257-271.
- CLAYTON, M. C., & HAYES, L. J. (1999). Conceptual differences in the analysis of stimulus equivalence. *The Psychological Record*, 49, 145-161.
- DUBE, W. V., GREEN, G., & SERNA, R. W. (1992). Auditory successive conditional discrimination and auditory stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior*, 59, 103-114.
- DYMOND, S., & BARNES, D. (1994). A transfer of self-discrimination response functions through equivalence relations. *Journal of the Experimental Analysis of Behavior*, 62, 251-267.
- DYMOND, S., & BARNES, D. (1995). A transformation of self-discrimination functions in accordance with the derived stimulus relations of Sameness, More than, and Less than. *Journal of the Experimental Analysis of Behavior*, 64, 163-184.
- DYMOND, S., & BARNES, D. (1996). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness and opposition. *The Psychological Record*, 46, 271-300.
- FIELDS, L., & VERHAVE, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior*, 48, 317-332.
- 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.
- HAYES, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 19-40). 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. 9-30). 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, 235-270.

- HAYES, S. C., & HAYES, L. J. (1989). The verbal action of the listener as a basis for rule-governance. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 153-190). New York: Plenum Press.
- HEALY, O., BARNES, D., & SMEETS, P. (1998). Derived relational responding as operant: The effects of between-session feedback. *The Psychological Record*, 48, 511-536.
- HOLTH, P., & ARNTZEN, E. (2000). Reaction times and the emergence of class consistent responding: A case for precurrent responding? *The Psychological Record*, 50, 305-337.
- KENNEDY, C. H. (1991). Equivalence class formation influenced by the number of nodes separating stimuli. *Behavioral Processes*, 24, 219-245.
- PILGRIM, C., & GALIZIO, M. (1996). Stimulus equivalence: A class of correlations or a correlation of classes? In T. R. Zentall & P. M. Smeets (Eds.), *Stimulus class formation in humans and animals* (pp. 173-195). B. V.: Elsevier Science.
- REESE, H. W. (1968). *The perception of stimulus relations: Discrimination learning and transposition*. New York: Academic Press.
- ROCHE, B., & BARNES, D. (1996). Arbitrarily applicable relational responding and sexual categorization: A critical test of the derived difference relation. *The Psychological Record*, 46, 489-507.
- ROCHE, B., & BARNES, D. (1997). A transformation of respondently conditioned stimulus functions in accordance with arbitrarily applicable relations. *Journal of the Experimental Analysis of Behavior*, 67, 275-301.
- ROCHE, B., BARNES-HOLMES, D., SMEETS, P. M., BARNES-HOLMES, Y., & MCGEADY, S. (2000). Contextual control over the derived transformation of discriminative and sexual arousal functions. *The Psychological Record*, 50, 267-291.
- ROCHE, B., STEWART, I., & BARNES-HOLMES, D. (1999). PsyScope: An easy-to-use graphical system for designing and controlling equivalence experiments. *The Experimental Analysis of Human Behavior Bulletin*, 17, 5-7.
- SIDMAN, M. (1971). Reading and auditory-visual equivalence. *Journal of Speech and Reading Research*, 14, 5-13.
- SIDMAN, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 93-114). Hove, England: Erlbaum.
- SIDMAN, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Author's Cooperative.
- SIDMAN, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, 74, 127-146.
- SPENCER, T. J., & CHASE, P. N. (1996). Speed analyses of stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, 65, 643-659.
- STEELE, D., & HAYES, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior*, 56, 519-555.
- STEWART, I., BARNES-HOLMES, D., ROCHE, B., & SMEETS, P. (2001). Generating derived relational networks via the abstraction of common physical properties: A possible model of analogical reasoning. *The Psychological Record*, 51, 381-408.
- WULFERT, E., & HAYES, S. C. (1988). The transfer of conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior*, 50, 125-144.

