

# Establishing Relational Responding in Accordance with More-than and Less-than as Generalized Operant Behavior in Young Children

Yvonne Barnes-Holmes<sup>1\*</sup>, Dermot Barnes-Holmes<sup>1</sup>, Paul M. Smeets<sup>2</sup>, Paul Strand<sup>3</sup>, and Patrick Friman<sup>4</sup>

<sup>1</sup>National University of Ireland, Maynooth, Ireland; <sup>2</sup>Leiden University, Holland;

<sup>3</sup>Washington State University, USA; <sup>4</sup>Girls and Boys Town Outpatient Behavioral Pediatrics and Family Services, USA

## ABSTRACT

The current study constitutes the first attempt to generate repertoires of relational responding, as generalized operant behaviors, when they are found to be absent in young children, using interventions suggested by Relational Frame Theory. Three children, aged between 4 and 6 years, were exposed to a basic problem-solving task that involved two or three identically-sized paper coins in an attempt to test and train patterns of relational responding in accordance with more-than and less-than. On each trial, the experimenter described how the coins compared to one another in terms of their value, and the child was then asked to pick the coin that would "buy as many sweets as possible". All three participants failed to pass baseline tests for specific patterns of arbitrary more and less responding. Interventions suggested by Relational Frame Theory, including training and testing across stimulus sets, were then successfully used to establish increasingly complex patterns of relational responding in all three children. Generalization tests demonstrated that the relational responding successfully generalized to novel stimuli and to a novel experimenter. In addition, the use of a non-contingent reinforcement condition for one participant, during which no improvement was made, together with contingency reversals for all children, indicated that the trained and tested relational responding may be considered a form of generalized operant behavior. These findings lend positive support to Relational Frame Theory's approach to derived relational responding, and to the functional analysis of human language and cognition. Alternative interpretations of the data are also considered.

*Key words:* More and less responding, generalized operant behavior.

## RESUMEN

*Establecimiento en niños pequeños de comportamiento relacional de acuerdo a más-qué y menos-qué como conducta operante generalizada.* El presente estudio constituye el primer intento de generar repertorios de responder relacionalmente como conducta operante generalizada cuando se encuentran ausentes en niños pequeños, siguiendo intervenciones sugeridas por la Teoría de los Marcos Relacionales.

\*Reprints may be obtained from the author: Department of Psychology, National University of Ireland, Maynooth, Maynooth, Co. Kildare, Ireland. E-mail: Yvonne.Barnes-Holmes@may.ie. This work was conducted as part of a Ph.D. research program undertaken by the first author, funded by the National University of Ireland, Maynooth.

Tres niños con edades entre 4 y 6 años fueron expuestos a una tarea básica de solución de problemas que implicó 2 o 3 billetes de idéntico tamaño, en un intento por probar y entrenar patrones de respuesta relacional de acuerdo a un patrón más-que y menos-que. En cada ensayo el experimentador describía comparar los billetes uno a otro en términos de su valor, y pedía al niño que tomara el billete con el que podría “comprar el mayor número posible de dulces”. Todos los participantes fallaron en pasar los tests de línea base para patrones específicos de responder arbitrario más y menos. Las intervenciones sugeridas por la Teoría de los Marcos Relacionales que incluyen entrenamiento y prueba a través de series de estímulos, se emplearon exitosamente para establecer patrones cada vez más complejos de responder relacional en los tres niños. Los tests de generalización demostraron que el responder relacional se generalizó exitosamente a estímulos nuevos y a un experimentador nuevo. Además, el uso de una condición de reforzamiento no-contingente con un participante durante la cual no se logró ninguna mejora, junto con reversiones de la contingencia para todos los niños, indicaron que el responder relacional entrenado y probado puede ser considerado una forma de conducta operante generalizada. Estos hallazgos proporcionan apoyo para la aproximación de la Teoría de los Marcos Relacionales al responder relacional derivado, y para el análisis funcional del lenguaje humano y la cognición. Interpretaciones alternativas de los datos son también consideradas.

*Palabras clave:* responder a más y menos, conducta operante generalizada.

Relational Frame Theory (RFT) has so far generated a range of studies that could all be described loosely as demonstration research (for a comprehensive review of this literature, see Hayes, Barnes-Holmes, & Roche, 2001). Specifically, some of these studies developed experimental procedures for demonstrating complex patterns of derived relational responding in human adult participants (e.g., Dymond & Barnes, 1995, 1996; Roche & Barnes, 1996, 1997; Roche, Barnes-Holmes, Smeets, Barnes-Holmes, & McGeady, 2000; Steele & Hayes, 1991; Wulfert & Hayes, 1988). More recently, one study conducted with adults (Healy, Barnes-Holmes, & Smeets, 2000), and two studies conducted with young children (Barnes-Holmes, Barnes-Holmes, Roche, & Smeets, 2001a, 2001b) have demonstrated that particular patterns of derived relational responding may demonstrate properties of generalized operant behavior. The latter three studies in particular provide evidence to support the argument that relational frames constitute generalized operant response classes.

In the first experiment reported by Healy et al. (2000), participants were initially exposed to conditional discrimination training on four matching-to-sample trial-types (A1-B1, B1-C1, A2-B2, and B2-C2). The participants were then tested for the formation of four combinatorially entailed derived relations (A1-C1, C1-A1, A2-C2, and C2-A2). Following exposure to this first cycle of training and testing, accurate or inaccurate feedback was delivered for the test performances. The next cycle of training and testing then began, but with a novel set of stimuli. This cycle of training and testing, using novel sets of stimuli for each cycle, continued until a participant responded in accordance

with the feedback across three consecutive stimulus sets. Once this stability criterion was reached, the type of feedback switched (accurate to inaccurate or vice versa) and the training and testing continued, using novel stimulus sets, until the performance again reached the stability criterion. In short, the feedback delivered for relational responding with earlier stimulus sets generalized to subsequent novel sets. This type of generalized consequential control across stimulus sets allowed the researchers to conclude that they had clearly demonstrated that derived relational responding can show one of the key properties of generalized operant behavior.

Nevertheless, Healy, et al. acknowledged that the relational repertoires that were targeted in the research were almost certainly established prior to the experiment (i.e., one would assume that undergraduates were capable of equivalence responding before entering the study). Consequently, the experimental procedures clearly influenced the participants' existing relational repertoires, but no evidence was provided that these repertoires were actually established during the experiment. In the words of Healy et al. (2000), "...the feedback influenced preexisting repertoires of generalized operant behavior, and did not establish those repertoires ab initio. Consequently, the current data do not provide strong evidence for the RFT view that derived relational responding is established, in the first instance, as generalized operant behavior" (pp. 224-225). Similar conclusions were also drawn by Barnes-Holmes, et al. (2001a, 2001b) in their work on derived symmetry with young children. The research reported in this article was designed to supplement the empirical work conducted within the conceptual framework of RFT, with a particular focus on the history that gives rise to what is referred to as derived relational responding (or relational framing) in normally-developing children.

In the current study, particular patterns of relational responding in accordance with the relational frame of more-than and less-than that appeared to be absent for a number of young children were selected. An attempt was then made to establish and manipulate those patterns as generalized operant behavior. Two recurrent themes run throughout this empirical work. First, in order to model the history of natural language interactions that supposedly give rise to relational framing in children, the experimental procedures involved relatively naturalistic speaker-listener exchanges. Second, we investigated the extent to which training across numerous sets of stimuli, and other techniques suggested by RFT, could be used to establish or facilitate specific patterns of relational framing in the children.

To test and train responding in accordance with more-than and less-than, a problem-solving task was designed that involved presenting a child with two or three identically-sized paper coins. On each trial, the experimenter described how the coins compared to one another in terms of their value, and the child was then asked to pick the coin that would "buy as many sweets as possible". On some trials involving two coins, for example, the child was told that one coin (coin A) would buy more sweets than another coin (coin B) (i.e., denoted as  $A > B$ ). On other trials involving three coins, for instance, the child was told that one coin would buy less sweets than a second coin, and that the second coin would buy less sweets than a third coin (i.e.,  $A < B < C$ ). Numerous sets of coins were used to test and train these types of relational performances. Each trial-type was designed to examine a particular pattern of transformation of function in accordance

with combinatorially entailed more-than or less-than relations. Imagine, for example, that on a particular trial, more-than relations (Crels) were established from coin A to coin B, and from coin B to coin C (i.e.,  $A > B > C$ ). If a participant then chooses coin A in the context of greatest value (Cfunc), according to RFT the function of coin A has been transformed in accordance with the combinatorially entailed more-than relations. In other words, the Crels and Cfunc transform the functions of coin A such that it is chosen over coin B and coin C.

With the presentation of three coins, four trial-types were constructed using this general approach, two that specified more-than relations (i.e.,  $A > B > C$  and  $C > B > A$ ), and two that specified less-than relations ( $A < B < C$  and  $C < B < A$ ). Of course, a total of 12 trial-types could be constructed using combinations of more-than and less-than relations among three coins presented in a linear sequence (e.g.,  $B > A > C$ ;  $C > A > B$ ;  $A < C < B$ ; etc.). Furthermore, a total of 24 trial-types could be constructed if nonlinear sequences are employed (e.g.,  $A > B < C$ ). The purpose of the current study, however, was not to conduct an exhaustive examination of the relational frame of comparison (although this has been done elsewhere, see Vitale, 2004), but rather to establish and manipulate specific patterns of relational responding that appear to be important constituent elements of this relational frame. The ultimate aim of the experiment was to establish for each child specific patterns of contextually controlled relational responding that would generalize to increasingly novel contexts (e.g., novel objects, settings, and a novel experimenter). It is important to emphasize from the outset, that the current study was not directly concerned with the analysis of sequence classes, order relations, or transitive inference (see Green, Stromer, & Mackay, 1993), although there may be functional overlap between these phenomena and the performances reported herein (see Discussion).

At this point, we would like to make clear our rationale in conducting and reporting this type of research. The current study employs a behavioral theory (RFT), grounded largely in basic research, to guide attempts to establish particular behavioral repertoires in young children. As such, we see the current research as making a contribution towards building bridges between basic and applied research domains. We clearly acknowledge that many of the interventions reported here are similar to those reported previously in the applied literature, and we are not attempting to suggest otherwise. Relational Frame Theory is a behavioral theory insofar as it draws on a range of different behavioral principles and procedures to explain complex human behavior (see Barnes-Holmes, Dymond, Roche, & Grey, 1999; Hayes, 1996; Hayes, et al., 2001). The overarching aim of the current research, therefore, is to determine whether RFT, as a basic behavioral theory, can successfully guide efforts to establish and manipulate specific patterns of relational responding in young normally-developing children.

## METHOD

### *Participants*

Three children (Participant 1, Participant 2, and Participant 3) participated in the current study. At the beginning of the experiment, P1, female, was 5 years and 5

months old; P2, male, was 5 years and 10 months old; and P3, male, was 4 years and 2 months old. At the end of the experiment, P1 was 5 years and 8 months old; P2 was 6 years and 1 month old; and P3, was 4 years and 7 months old. All three children attended day-care facilities in Ireland, the first two participants were enrolled in a crèche in Cork, and the third participant was enrolled in a crèche in Dublin. The children were chosen on the basis of parental consent, and that neither their parents nor their crèche supervisor had identified them as presenting a learning difficulty.

### *Setting and Materials*

Each session was conducted in a quiet room free from distraction. The children participated individually. The experimenter and child sat side-by-side at a small wooden table during most of the sessions. During generalization tests (described later), a novel experimenter and the child sat together on the floor. Forty-five identically-sized colored paper coins were employed throughout the study. These were described to the children as “coins”, and this label will be used throughout the current paper. There were fifteen blue coins, fifteen red coins, and fifteen green coins, and each coin was marked with a different pattern (i.e., no two coins were identical). The coins were divided into fifteen sets, each containing three coins -- one blue, one red, and one green (referred to as Sets 1-15). The three coins contained in each set were designated as A, B, and C (participants never saw these labels). Across sets, coins were randomly designated as A, B, or C and thus control by either the dimension of color or pattern was eliminated. Only one set of coins was used at any one time.

Each set was placed on a background of white A4 paper (referred to as the stimulus sheet). During some trials, the stimulus sheet contained one or two black printed arrows, with each arrow positioned between each pair of coins arranged horizontally. Each arrow pointed either to the right or to the left. When three coins were presented simultaneously, the stimulus sheet contained two identical arrows, both pointing in the same direction. For example, when coins A, B, and C were presented, one arrow was presented between coins A and B, and another arrow between coins B and C. The words “BUY MORE” or “BUY LESS” were printed above each arrow, and in cases of two arrows, both were accompanied by the same words, (i.e., both had “BUY MORE” or “BUY LESS” above the arrow). In subsequent phases of the study, the arrows and the text were systematically removed (because none of the children could read these words). Stimulus sheets containing only the text (i.e., with arrows removed), and blank stimulus sheets containing neither arrows nor text were subsequently used for this purpose.

A number of additional sets of stimuli was employed throughout the study to test for generalization. These included: books, compact disc (CD) covers, drinking glasses, pencils, and spoons. All of the generalization objects in each category (e.g., books) were identical in size. Other materials were employed as reinforcers including colored beads, commercially available children’s stickers and sweets. The reinforcers and an upright glass jar were placed on a wooden tray. The tray was placed to the left, and slightly in front of the experimenter throughout each session.

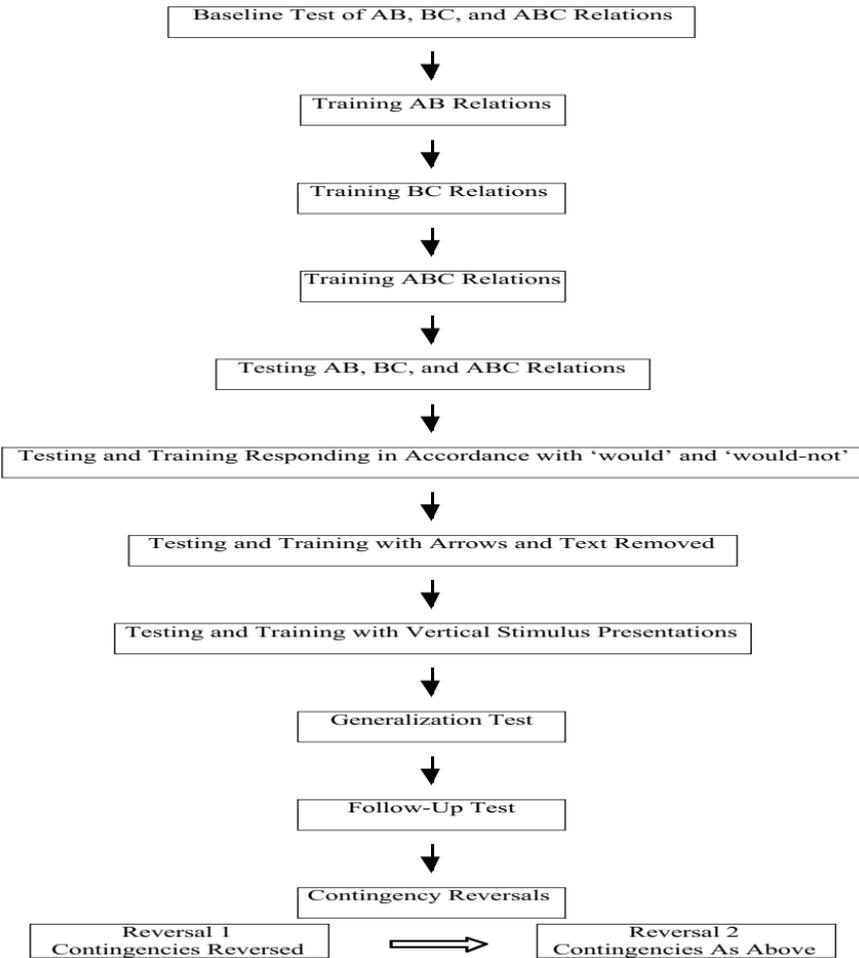


Figure 1. A schematic representation of the experimental sequence.

*Programmed Consequences*

A correct response consisted of the child pointing to the correct coin, and was followed by the words “Yes, you are correct. Good girl/boy. Take a bead.” An incorrect response was defined as making an incorrect choice or emitting no response within 10 seconds of the instruction. After collecting eight beads in the glass jar, the child was allowed to select a sticker/sweet from the wooden tray. Punishment during training trials consisted of the experimenter saying: “No, this is not correct. You lose a bead.” In this case, the experimenter removed a bead from the jar and placed it back in the tray, and the next training trial began. If a child made any comment during a trial, the experimenter simply replied “We can talk after we have finished our work.” No

programmed consequences followed any test trial.

*General Procedure*

A schematic representation of the experimental sequence is provided in Figure 1. All three participants were exposed to this basic experimental sequence, although the actual training and testing to which each child was exposed depended upon his or her performances on the various tests. Testing and training trials were identical, except that the latter were each consequted with corrective feedback, whereas the former were not. Testing and training trials were always presented in blocks of eight trials.

Participants were first exposed to a baseline test to determine whether they could respond in accordance with more-than and less-than relations among the three coins in Set 1. This test consisted of twelve trial-types, and these are depicted in Figure 2. There were four trial-types involving AB relations and four trial-types involving BC relations, all of which involved the presentation of only two coins. There were also four trial-types involving ABC relations, during which all three coins were presented. All of these trial-types are described in detail below. The baseline test involved two exposures to each of the twelve trial-types, presented in the order of eight AB trials, followed by eight BC trials, and finally eight ABC trials. In each test or training trial, the child was required to point to a particular coin. Pointing to two or more coins, even if one of the

| AB Relations        | BC Relations        | ABC Relations       |                     |
|---------------------|---------------------|---------------------|---------------------|
| LESS THAN<br>A → B* | LESS THAN<br>B → C* | LESS THAN<br>A → B  | LESS THAN<br>B → C* |
| LESS THAN<br>A* ← B | LESS THAN<br>B* ← C | LESS THAN<br>A* ← B | LESS THAN<br>B ← C  |
| MORE THAN<br>A* → B | MORE THAN<br>B* → C | MORE THAN<br>A* → B | MORE THAN<br>B → C  |
| MORE THAN<br>A ← B* | MORE THAN<br>B ← C* | MORE THAN<br>A ← B  | MORE THAN<br>B ← C* |

\* Indicates correct choice.

→ Arrow indicates direction in which Experimenter pointed.

*Figure 2.* Trial-types used for testing and training the relations of more-than and less-than. In trials involving the presentation of text only, the arrows were removed and the participant was exposed to text only above the coins. In trials involving the vertical presentation of coins, the participant saw neither arrows nor text, only the coins positioned on a blank sheet of A5 paper, and Coin A was always positioned above Coin B, which was always positioned above Coin C. In the generalization test, trial-types were identical to those presented in Figure 2 and used throughout the experiment, but the stimuli were either CD covers, books, pencils, drinking glasses or spoons randomly presented on the floor of the experimental room.

choices was correct, was recorded as an incorrect response. In order to pass the baseline test, participants were required to achieve a minimum of 21 correct responses (out of 24) overall without producing two or more incorrect responses on any one relation.

In general, sessions lasted no more than 20 minutes per day, and the children were exposed to a maximum of four sessions per week. When sessions lasted more than 10 minutes, a break of 5 minutes was provided mid-way through the session. At the beginning of each block of training or testing trials, the experimenter always asked the child "Do you want to do some more work?" If the child indicated that s/he did want to do more, the experimenter continued as planned. If, however, the child responded negatively (or indicated during a training or test block that s/he wished to stop), the experiment was terminated for that day. If the child had reached a training criterion or passed a test during the previous block, in the next session the experimenter continued with the next planned stage of the experiment. If, however, the participant had failed to reach a training criterion or pass a test during the previous block (or asked to stop at any point during a block) the next planned stage was not presented. Instead, the next training or test block normally involved some form of reduction in the complexity of the previously presented stage (e.g., presenting two coins rather than three coins).

*Training AB relations.* After the baseline test (on which all three participants failed), each child received explicit training on the AB relations in blocks of eight trials, involving only coins A and B. Coin A was consistently placed on the left-hand side of the stimulus sheet, with Coin B on the right. On the first trial of each session, the experimenter placed the bead container on the table and positioned the coins according to the appropriate trial-type. The participant was first told:

"We are going to play a birthday game." The following instructions were then given: "I want you to imagine that it is your birthday today, and you have to go to the shops to get sweets for your birthday party. If I tell you that this coin (e.g., experimenter pointed to coin A) buys less (or more) sweets than this coin (experimenter pointed to coin B), which would you take to buy as many sweets as possible?"

On subsequent trials, shorter instructions were provided as follows "If this coin (e.g., experimenter pointed to coin A) buys more (or less) sweets than this coin (experimenter pointed to coin B), which would you take to buy as many sweets as possible?"

There were four trial-types in each block of AB relations (two less-than trial-types, and two more-than trial-types), and each type was presented twice in a random order without replacement (see Figure 2). Each of these trial-types may be described as follows: A buys less (sweets) than B; B buys less than A; A buys more than B; and B buys more than A. Two of the trial-types involved the experimenter pointing to the A coin first (e.g., A buys less than B), whereas the other two trial-types involved the experimenter pointing to the B coin first (e.g., B buys more than A), depending on the relation being stipulated. Participants were required to reach a mastery criterion of eight consecutively correct responses on the AB training trials before proceeding immediately to training on the BC relations.

*Training BC relations.* The procedure for training the BC trial-types (see Figure 2) was identical to that employed with the AB trial-types, except that coins B and C were used instead of A and B. Participants were required to reach the mastery criterion before proceeding to training on the ABC relations.

*Training ABC relations.* The four ABC trial-types involved the use of all three coins, A, B, and C. The coins were positioned from left to right in the order of A, B, and C, respectively (the coins remained in these positions throughout each block of trials). An example of the instructions provided during ABC training is as follows: "If this coin (e.g., experimenter pointed to coin A) buys less sweets than this coin (experimenter pointed to coin B), and if this coin (experimenter pointed again to coin B) buys less sweets than this coin (experimenter pointed to coin C), which would you take to buy as many sweets as possible?" Across trials, the experimenter pointed to the A coin first or to the C coin first, depending on the trial-type. The four trial-types were as follows: A buys less than B which buys less than C; C buys less than B which buys less than A; A buys more than B, which buys more than C; and C buys more than B which buys more than A (see Figure 2). Each trial-type was presented twice in a random order without replacement in a block of eight training trials.

*Testing AB, BC, and ABC relations.* This test consisted of 24 test trials and was identical to the baseline test to which participants had been exposed initially. Successful training on AB, BC, and ABC relations was always followed by a test of all three relations using a new set of coins. The 24 test trials were divided up into three blocks of eight trials, eight AB trials, eight BC trials, and eight ABC trials, identical to those described above. The 24 test trials were also presented in this order (i.e., AB, BC, and ABC, respectively). In order to pass the test, participants were required to achieve a minimum of 21 correct responses out of 24, without producing two or more incorrect responses on any one relation (e.g., if they responded incorrectly more than once to the 'A is worth more than B' relation, the test performance was recorded as a fail). It is important to emphasize that training and testing were identical except for the provision of programmed consequences and the use of novel stimulus sets. At the beginning of each test, participants were given the same instructions as during training, and were also informed "This time I can't tell you whether you are right or wrong."

*Responding in accordance with 'would' and 'would-not'.* When participants successfully passed the full test of all three types of relations (i.e., AB, BC, and ABC), they were exposed to a similar test involving 'would' and 'would-not' trial-types. From an RFT perspective, the phrases "more-than" and "less-than" functioned as Crels for the relational frame of comparison, whereas "would" and "would-not" functioned as Crels for the relational frame of distinction (in this case logical not). Responding to these tasks, therefore, was under the synergistic control of two separate Crels, thereby introducing additional flexibility at the level of contextual control into the children's relational performances. On these trials, participants were instructed as follows: "This time, I will sometimes ask which would you take to buy as many sweets as possible, and other times I will ask which would you not take to buy as many sweets as possible?" Participants were required to indicate which coin they *would* or *would not* select in order to buy as many sweets as possible. All participants were first exposed to the

would/would-not trials as test trials. In the full 24-trial test presented during baseline and then again after the explicit training described above, there were two exposures to each of 12 trial-types (four AB trials, four BC trials, and four ABC trials), all of which required the participant to select the coin that s/he *would* choose in order to buy as many sweets as possible. In order to incorporate would-not trials, this test was modified such that there was now only one exposure to each of the 12 trial-types as would trials, whereas in the second exposure to these trial-types they were now presented as would-not trials. In other words, there were eight AB trials (four would and four would-not), eight BC trials (four would and four would-not), and eight ABC trials (four would and four would-not), respectively. The would and would-not trials were presented randomly within each section of the test. For illustrative purposes, consider the following example. In the block of eight AB trials, two of the trial-types may be summarized as follows: (i) A buys more sweets than B: Which coin *would* you choose? and (ii) A buys more sweets than B: Which coin *would you not* choose? The same format was applied to the blocks of BC and ABC trial-types. All participants were first tested on would/would-not trial-types. If a participant failed this test, he or she was then exposed to the same trial-types with feedback until reaching the criterion of eight consecutively correct responses on the AB, BC, and ABC trials, respectively. Participants were thereafter reexposed to the full would/would-not test (without feedback) using a novel stimulus set.

*Removing the arrows and text.* After participants had demonstrated 'would' and 'would-not' responding, specific features of the stimulus presentation were systematically altered, in order to eliminate the arrows and/or text as possible sources of stimulus control. First, participants were exposed to the would/would-not test as above but the arrows located between each pair of coins were removed, although the text that had been positioned above the arrows was retained. If participants failed this test, they received explicit training of the trials (i.e., with feedback) presented in this format until they passed a test (without feedback) on a new set. When participants passed the test, they were reexposed to the test with the text that had been positioned above the arrows now removed. That is, at this point only the coins were presented on the stimulus sheets with both arrows and text removed. Once again, participants were first exposed to this modification in the form of test trials and explicit training was only provided where appropriate. Where participants were provided with explicit training in the absence of the text, this was followed by test trials involving the same stimulus presentation using a novel stimulus set. In all subsequent trials for each participant throughout the experiment, the arrows and text were never reintroduced.

*Vertical stimulus presentations.* All of the previously described trial-types involved the horizontal presentation of coins placed adjacent to one another (i.e., A beside B, B beside C, or A beside B beside C). After participants had successfully completed all of the training and testing outlined above with stimuli presented in this manner, they were exposed to a test in which the orientation of the stimuli was changed from horizontal to vertical. This alteration in the stimulus presentation permitted the elimination of the horizontal presentation as a possible source of stimulus control. For example, instead of stimuli being placed with A, B, and C from left to right, respectively, the coins were

now positioned A, B, and C vertically, with A at the top, B in the middle, and C at the bottom of the stimulus sheet. Once again, explicit training on a vertical stimulus presentation was provided only where appropriate and was followed by a test on a novel stimulus set. Once a participant had passed the test involving the vertical presentation of coins, this type of presentation was never used again with that participant.

*Generalization test.* When participants had completed all of the training and test procedures outlined above, they were exposed to a generalization test. This test involved a novel set of three identically-sized objects randomly positioned around the floor of the experimental room with the stimulus sheet removed. This generalization test contained identical trial-types to the would/would-not test, except that they involved other objects instead of coins. A novel experimenter conducted all generalization tests. The novel experimenter was provided with an appropriate script of the relevant question to be asked on each trial at the beginning of each session. The novel experimenter was not required to record responses (this was done by the original experimenter). The novel experimenter was also explicitly instructed not to deduce the correct answer to each trial because doing so might interfere with the experiment.

*Follow-up tests.* Follow-up tests, where possible, were conducted one month after the completion of testing and training to determine if the relational performances remained intact across extended periods of time (see Rehfeldt & Hayes, 2000; Saunders, Wachter, & Spradlin, 1988). These tests involved: a novel set of coins; the horizontal presentation of stimuli, with no arrows or text; would and would-not trial-types; and no stimulus sheet.

*Contingency reversals.* When participants had passed all of the tests outlined above, the reinforcement contingencies were reversed (i.e., Reversal 1) in order to determine the operant nature of the performances that had been demonstrated. In effect, participants were now required to respond away from the coin the choice of which would have been reinforced previously. For example, given the relation 'A worth more than B', selecting B was now reinforced, whereas selecting A was reinforced previously. Reversal 1, therefore, involved reexposing participants to all of the procedures described above but with the reinforcement contingencies reversed (see Figure 1). After participants had passed all of the tests contained in Reversal 1, including the generalization test, a second reversal (i.e., Reversal 2) was introduced in order to complete an A-B-A reversal design for each participant. In Reversal 2, the contingencies were reversed a second time, and the original reinforcement contingencies were reinstated. That is, given the relation 'A worth more than B', for example, selecting coin A was reinforced as before. Once again, therefore, Reversal 2 involved exposing participants to all of the procedures described above now for a third time, but with the reinforcement contingencies as in the original set of exposures.

### *Inter-observer reliability*

Approximately 25 percent of training and testing trials were observed by an independent observer, who had no knowledge of experimental psychology. The observer could not see the experimenter's data sheet during the experimental sessions. The

Table 1. Sequence of training and testing, number of training trials, and test outcomes for Participant 1 on the relations of more-than and less-than.

| Condition    | Training/Test |         | Stimulus |  | No. of Training Trials/<br>Test Outcomes |
|--------------|---------------|---------|----------|--|--|
|              | Exposures     | Session | Set      | Training/Test Type                             |  |
| Baseline     | 6 Tests       | 1-6     | 1-3      | Would  | FFFFFF                                   |
| Intervention | Training      | 7-11    | 3        | Would  | 136                                      |
|              | 1 Test        | 11      | 4        | Would  | F  |
|              | Training      | 12      | 4        | Would (ABC only)                               | 14                                       |
|              | 2 Tests       | 12-13   | 5        | Would; Would/Not                               | P;F                                      |
|              | Training      | 14      | 5        | Would/Not                                      | 28                                       |
|              | 1 Test        | 14      | 6        | Would/Not                                      | F  |
|              | Training      | 15-16   | 6        | Would/Not                                      | 64                                       |
|              | 7 Tests       | 16-22   | 7-8      | Would/Not; Text Only; Vertical;<br>Gen. (CD's) | FP; FP; PP<br>P                          |
| Follow-Up    | 2 Tests       | 23-24   | 9        | Would/Not                                      | FP                                       |
| Reversal 1   | Training      | 25-26   | 9        | Would  | 52                                       |
|              | 1 Test        | 27      | 10       | Would  | F  |
|              | Training      | 28      | 10       | Would  | 24                                       |
|              | 1 Test        | 28      | 11       | Would  | F  |
|              | Training      | 29-30   | 11       | Would  | 69                                       |
|              | 3 Tests       | 30-32   | 12       | Would; Would/Not; Gen. (Books)                 | P;P;P                                    |
| Reversal 2   | Training      | 33      | 12       | Would  | 28                                       |
|              | 1 Test        | 33      | 13       | Would  | F  |
|              | Training      | 34      | 13       | Would  | 25                                       |
|              | 1 Test        | 34      | 14       | Would  | F  |
|              | Training      | 35      | 14       | Would  | 25                                       |
|              | 4 Tests       | 35      | 15       | Would; Would/Not; Gen. (Pencils)               | FP;P;P                                   |

P = Pass; F = Fail: Reading from left to right

FP indicates that the participant failed the first exposure to a test, and passed the second exposure to the same test.

Gen. = Generalization test.

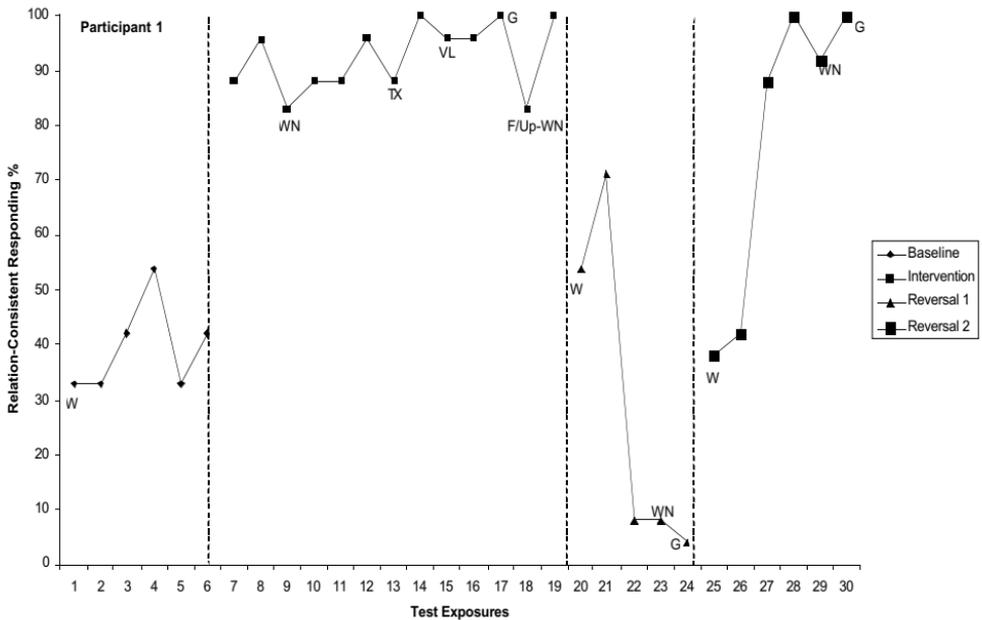
observer and experimenter disagreed on a total of only four training trials and two test trials.

#### PROCEDURE AND RESULTS

Given the nature of the study, the procedural details pertaining to each participant will be described in the context of the results. The complete procedure and results for P1 will be presented, but for P2 and P3 only those features of the experiment that differ from P1 will be described (i.e., the entire experimental sequence for P2 and P3 will not be outlined).

#### *Participant 1*

The training and test data for P1 are outlined in Table 1, and the participant's performance during each test exposure is presented in Figure 3. Participant 1 was first



**Figure 3.** Percentage of relation-consistent responses for Participant 1 on tests of the relations of more-than and less-than. The participant's responses during training are not depicted. Letters adjacent to data points indicate the type of stimulus presentation or experimental phase in operation. Data points that are not accompanied by letters involve the stimulus presentation or condition indicated by the previously marked data point.

(W = would responding; WN = would and would-not responding; TX = text-only stimulus presentation; VL = vertical stimulus presentation; G = generalization test; F/Up = follow-up).

exposed to six consecutive baseline tests of the AB, BC, and ABC relations using three sets of coins (i.e., Sets 1-3). Set 1 was employed in Sessions 1 and 2, Set 2 was employed in Sessions 3 and 4, and Set 3 was employed in Sessions 5 and 6. Participant 1 failed to pass any of the six tests, and produced a highest score of 13 out of 24, or 54% correct. Given the lack of improvement across testing, explicit training of the relations was introduced, beginning in Session 7 with the AB relations, using Set 3.

After 40 exposures to the training trials (i.e., 5 blocks of 8 trials), P1 was still failing to produce eight consecutively correct responses on the AB relations. At this point, therefore, a response-cost procedure was introduced in which every incorrect response was followed by the removal of one of the existing beads from the jar, and the participant was instructed to try again on the same trial. If a correct response was then emitted, the participant received verbal praise, but was not allowed to select a bead (this modification was based on the concept of the *learn unit* described in Greer, Phelan, & Sales, 1993). This modification was employed on all subsequent training sessions involving this participant and the two other participants in all subsequent training trials. After a further 37 training trials, P1 produced eight consecutively correct

responses on the AB relations (i.e., 77 AB training trials in total). As indicated in Table 1, P1 required a total of 136 training trials in order to complete training on the AB, BC, and ABC relations across Sessions 7-11 (i.e., she required 25 and 34 training trials on the BC, and ABC relations, respectively). This training was followed by a test of all three relations using a new set of coins (i.e., Set 4). Participant 1 now produced 21 out of 24 correct responses, a dramatic improvement from baseline test performances. Although this test performance constituted a substantive improvement, all three errors produced by the participant occurred on the ABC less-than relations, and so the participant was retrained on the ABC relations only. When she reached eight consecutively correct responses on the ABC relations in 14 training trials, P1 was reexposed to a complete test of all relations using a new set (i.e., Set 5). In Session 12, she produced 23 out of 24 correct responses, which constituted the first successful test performance.

Having now passed the initial test of AB, BC, and ABC relations, the would/would-not test was introduced. Participant 1 produced 20 out of 24 correct responses on this test (Session 13), and although this performance was high, it did not constitute a pass. She was subsequently trained explicitly to respond appropriately to the would and would-not trial-types. Twenty-eight training trials were required for her to reach the mastery criterion on all three types of relations. Following this training, she was retested (in Session 14) on the would/would-not trial-types using a new set of coins (i.e., Set 6). Participant 1 produced 21 correct responses, with two errors occurring on the same relation, and so she was retrained on all relations using the same set of coins. This time she required 64 training trials to achieve the mastery criterion, and she was then (in Session 16) reexposed to the would/would-not test using a new set of coins (Set 7). She again produced 21 correct responses out of 24, with two errors on the same relation. Because improvement across blocks of test trials is a relatively common finding in the derived stimulus relations literature (e.g., Devany, Hayes, & Nelson, 1986), P1 was reexposed to the same test in the subsequent session without further training. In Session 17, the participant passed the would/would-not test with 23 correct responses.

Having now passed the initial test of all relations involving both would and would-not responding, features of the stimulus presentation were altered. (Until specified otherwise, all subsequent training and testing sequences incorporated would and would-not trial-types). In Session 18, the 'text only' test was introduced, in which all arrows on the stimulus sheet were removed, and only the text that had been positioned above the arrows was retained. Participant 1 produced 21 correct responses out of 24 and again, although this constituted a fail, she was retested without further training. She then produced perfect responding (i.e., 24 out of 24) on this test. In the following session, the text was removed, and the orientation of the stimuli was altered from horizontal to vertical. In the first test involving a vertical presentation (in Session 20), P1 passed with 23 correct responses. Because the same stimulus set had now been used for a number of tests, the participant was retested immediately using a novel set (i.e., Set 8), and she passed again.

Following this successful performance of would and would-not responding using the vertical presentation of stimuli, a generalization test was conducted using two and three identically-sized CD covers, randomly positioned around the floor of the experi-

mental room. As indicated previously, a novel experimenter was employed for this and for all subsequent generalization tests. During this test, P1 immediately produced perfect responding.

Approximately one month later (in Session 23), a follow-up test was conducted (with the original experimenter) involving a new set of coins, in order to determine whether the relational performances were retained in the child's repertoire. This test involved would and would-not responding, and the coins were presented in horizontal positions. All subsequent stimulus presentations were horizontal. P1 produced 20 correct responses (i.e., failing by one response). The child was then retested without further training and she produced a perfect test performance. From these data, it appeared that the arbitrary relations of more-than and less-than, involving responding in accordance with would and would-not, and incorporating some degree of generalization, had been established in the repertoire of P1.

In the next part of the study, P1 was exposed to Reversal 1 (see Figure 3), in which the baseline contingencies were reversed. The participant was now required to respond *away* from the coin, the choice of which would have been reinforced during baseline trials. Because it seemed uncertain whether this task would prove difficult for her, Reversal 1 training began with training on would-only trials. In Sessions 25-26, P1 required 52 training trials in order to reach the mastery criterion on all three types of relations during the contingency reversal. On a subsequent test (in Session 27) of these relations, involving a new set of coins (i.e., Set 10), she failed to pass the reversal test, producing 11 "reversal" responses out of 24 (i.e., 13 "correct" responses based on the prereversal contingencies). The participant was subsequently given further explicit training on the reversed relations and reached the mastery criterion in 24 training trials. On a subsequent test, the performance of the participant deteriorated, and she now produced only 6 reversed responses out of 24 (i.e., 18 previously "correct" responses). She was then (in Sessions 29-30) exposed to further explicit training on the reversed relations, and required a total of 69 training trials in order to reach the mastery criterion. Immediately after this training, she passed a test of the reversed relations by producing 22 out of 24 reversed responses (i.e., only 2 previously "correct" responses). Furthermore, in subsequent sessions (31-32), P1 passed, without further training, both a would/would-not test and a generalization test of the reversed relations (i.e., with books randomly positioned on the floor). From these data, as depicted in Figure 3, it is apparent that the original pattern of responding, observed before the introduction of the contingency reversal, had now been successfully modified.

Having passed all of the tests contained in the first reversal, P1 was now exposed to Reversal 2 (see Figure 3), in which the original reinforcement contingencies were reinstated. This reversal phase once again began with training on would-only trial-types. In Session 33, she reached the mastery criterion on all of the relations in only 28 training trials. In a subsequent test of these relations involving a new set of coins (i.e., Set 13), she failed the would-only test by producing only 9 correct responses out of 24 (i.e., 15 responses in accordance with the previously reversed reinforcement contingency). She was again explicitly trained on these relations, and reached the mastery criterion in only 25 training trials. Participant 1 was subsequently retested using a new

Table 2. Sequence of training and testing, number of training trials, and test outcomes for Participant 2 on the relations of more-than and less-than.

| Condition    | Training/Test |         | Stimulus |                                       | No. of Training Trials/<br>Test Outcomes |
|--------------|---------------|---------|----------|---------------------------------------|--|
|              | Exposures     | Session | Set      | Training/Test Type                    |  |
| Baseline     | 4 Tests       | 1-2     | 1-2      | Would                                 | FFFF                                     |
| Intervention | Training      | 3-5     | 2        | Would (AB only)                       | 104*                                     |
|              | 2 Tests       | 6       | 3        | Would Non-arbitrary; Arbitrary        | P;F                                      |
|              | Training      | 7-9     | 3        | Would Interpolated (AB only)          | 91                                       |
|              | 1 Test        | 9       | 4        | Would (AB only)                       | F  |
|              | Training      | 10      | 4        | Would                                 | 24                                       |
|              | 5 Tests       | 11-12   | 5        | Would; Would/Not; Text Only; Vertical | P;P;P;FF                                 |
|              | Training      | 13      | 5        | Vertical                              | 42                                       |
|              | 2 Tests       | 13-14   | 6        | Vertical; Gen. (Glasses)              | P;P                                      |
| Follow-Up    | 1 Test        | 15      | 7        | Would/Not                             | P  |
| Reversal 1   | Training      | 15-16   | 7        | Would                                 | 26                                       |
|              | 3 Tests       | 16-17   | 8        | Would; Would/Not; Gen. (Books)        | P;P;P                                    |
| Reversal 2   | Training      | 18      | 8        | Would                                 | 29                                       |
|              | 2 Tests       | 18      | 9        | Would; Would/Not                      | P;F                                      |
|              | Training      | 19      | 9        | Would/Not                             | 34                                       |
|              | 2 Tests       | 19      | 10       | Would/Not; Gen. (Pencils)             | P;P                                      |

P = Pass; F = Fail: Reading from left to right

\* Indicates that the participant failed to reach the mastery criterion during training.

Gen. = Generalization test.

set of coins (Set 14), but again failed the test by producing only 10 correct responses. In the subsequent session (35), she was retrained again, and reached the mastery criterion in 25 training trials. The participant was then immediately retested on a new set of coins (Set 15) and the test performance improved dramatically to 21 correct responses, but with two errors occurring on the same relation. She was immediately retested, and she produced a perfect test performance. In the same session, she was then tested on would and would-not trial-types using the same set of coins. In this test, P1 immediately passed with 22 correct responses. The final test was a generalization test using pencils randomly positioned on the floor of the experimental room. On this final test, P1 immediately produced a perfect test performance.

### Participant 2

The training and test data for P2 are outlined in Table 2, and the participant's performance during each test exposure is presented in Figure 4. The testing and training procedures employed with this participant were very similar to those employed with P1, except that a number of minor alterations were required, especially in establishing initial responding to the AB, BC, and ABC relations. Participant 1 had required a total of 150 training trials on the initial AB, BC, and ABC relations before passing the first baseline test. However, after 104 exposures to only the AB training trials (in Sessions

3-5), P2 still failed to reach the mastery criterion. At this point in P2's training, it seemed appropriate to turn our attention to nonarbitrary stimulus relations.

According to RFT, a history of reinforcement for responding in accordance with nonarbitrary relations (i.e., responding that is controlled by the formal properties of the stimuli) provides an important historical context for the establishment of their arbitrary counterparts in a child's behavioral repertoire. For RFT, nonarbitrary relational control (e.g., learning to pick the larger of two coins when asked to pick the bigger coin) is functionally distinct from arbitrary relational control (e.g., learning to pick the more valuable, but smaller, of two coins). For example, only in the most artificial of learning environments could one imagine a child responding in accordance with the arbitrary

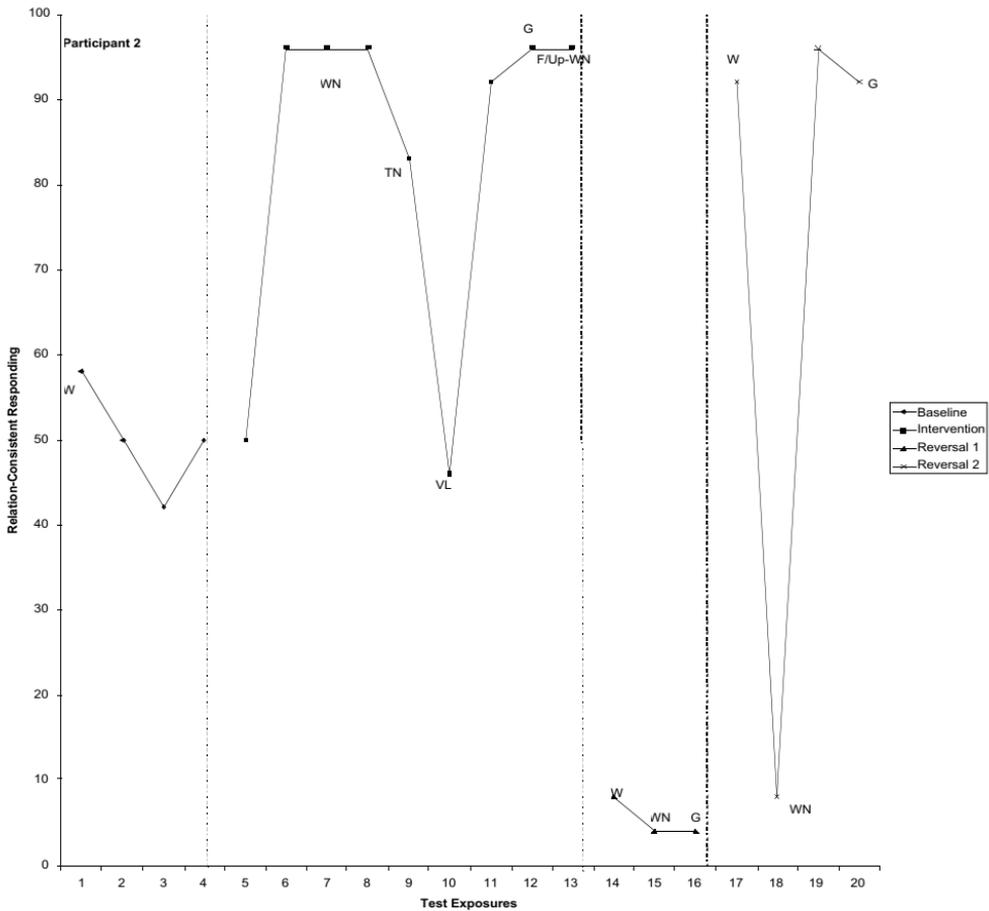


Figure 4. Percentage of relation-consistent responses for Participant 2 on tests of the relations of more-than and less-than.

relations of more-than and less-than before first demonstrating the nonarbitrary class of this relational responding (Hayes et al., 2001). In Session 6, therefore, a test of the nonarbitrary more-than and less-than relations that involved placing sweets on top of two coins was introduced with P 2. This test consisted of four trial-types that matched the arbitrary trial-types used previously (see Figure 2). Two trial-types consisted of two sweets placed on top of coin A, and one sweet placed on top of coin B (i.e., A was physically more than B, and B less than A). The two other trial-types consisted of the reverse arrangement with two sweets placed on top of coin B and one sweet on top of coin A (i.e., B was physically more than A, and A less than B). The location of the coins was also alternated randomly, such that on half of the test trials coin A was positioned on the left with coin B on the right, and on the other half of the trials coin B was positioned on the left with coin A on the right. The test therefore consisted of one block of eight trials, with each trial-type randomly presented twice, once in each location, without replacement. At the beginning of each trial the participant was simply asked "Which coin has more?" No feedback was provided during this test. Participant 2 produced perfect responding on this nonarbitrary test.

After passing the nonarbitrary test, the participant failed a subsequent exposure to a baseline arbitrary test. Because this participant had passed the nonarbitrary test, but failed the arbitrary tests, an intervention involving interpolating arbitrary and nonarbitrary training trials was introduced in an attempt to establish the mutually entailed arbitrary more-than and less-than relations. During Sessions 7 and 8, seven blocks of eight training trials of AB relations only were presented, with each block containing four arbitrary and four nonarbitrary trial-types, as described above. The first training trial was a nonarbitrary trial-type, and this was always followed by an arbitrary trial-type. All trial-types during this part of the training were presented in that order. After 56 training trials, P2 had still failed to produce eight consecutively correct responses on the interpolated arbitrary training trials (and showed no sign of improvement), but produced no errors on the nonarbitrary trials. At this point, simply interpolating nonarbitrary and arbitrary training trials was discontinued, and an alternative procedure was introduced to train the arbitrary relations between two coins.

In Session 9, nonarbitrary trials were used to correct errors conducted on arbitrary trials, without presenting any nonarbitrary trials independently. On an arbitrary training trial, for example, the participant was presented with two coins, with no sweets placed on top, and asked "If this coin (e.g., experimenter pointed to coin A) has more sweets than this coin (experimenter pointed to coin B), which would you take to buy as many sweets as possible"? If the child produced an incorrect response, the experimenter transformed the trial into a nonarbitrary trial-type by placing two sweets on top of the coin which was worth 'more', and one sweet on top of the coin which was worth 'less', according to the arbitrary relation specified. The participant was then asked "Now which has more?" During this type of training, the participant produced no errors on the nonarbitrary correction trials. A correct nonarbitrary (corrective) response was followed by verbal praise, but no bead could be selected. A new arbitrary training trial was then presented. After 35 training trials conducted in this way, P2 finally produced eight consecutively correct responses on the arbitrary AB relations. As indicated in Table 2,

he required 91 training trials in total with nonarbitrary interventions to reach the mastery criterion on the AB relations only. He was then immediately exposed to a block of eight test trials of arbitrary relations involving only two coins (A and B) from a novel set (i.e., Set 4), but failed. In Session 10, the participant was introduced once again to a block of eight arbitrary training trials of AB relations (i.e., without nonarbitrary interventions), and he produced eight consecutively correct responses immediately. In an effort to maintain a relatively high level of reinforcement for on-task behavior, training on the BC and ABC relations followed immediately. He completed this training without error (i.e., a total of 24 training trials were required to complete all three types of relations in Session 10). At this point, P2 had now successfully completed initial training of all three arbitrary relations for the first time, and had required a total of 219 training trials to reach the mastery criterion on all three baseline arbitrary relations. He passed the baseline arbitrary test for the first time in Session 11.

Unlike P1, P2 passed the would/would-not test without training, but after an absence of three weeks, he failed two exposures to the vertical test, and required explicit training of the relations presented vertically. He immediately passed a subsequent vertical test. In Reversals 1 and 2, he adjusted quickly to the altered contingency arrangements, and generally required very little explicit training. Unlike P1, he did not require repeated training exposures to the initial reversed relations involving would-only trial-types, although in Reversal 2, he did require some explicit training on the

*Table 3.* Sequence of training and testing, number of training trials, and test outcomes for Participant 3 on the relations of more-than and less-than.

| Condition                | Training/Test |         | Stimulus |  | No. of Training Trials/<br>Test Outcomes |
|--------------------------|---------------|---------|----------|--|--|
|                          | Exposures     | Session | Set      | Training/Test Type                             |  |
| Baseline 1<br>Non-Cont'g | 3 Tests       | 1-2     | 1-2      | Would  | FFF                                      |
| Reinforcement            | Training      | 3-7     | 2        | Would (AB only)                                | 224*                                     |
| Baseline 2               | 3 Tests       | 8-10    | 2-3      | Would  | FFF                                      |
| Intervention             | Training      | 11-17   | 3        | Would (AB only)                                | 121                                      |
|                          | Training      | 18-20   | 3        | Would (BC only)                                | 75                                       |
|                          | Training      | 21-23   | 3        | Would (ABC only)                               | 40*                                      |
|                          | Training      | 24      | 3        | Would (AB only)                                | 18                                       |
|                          | Training      | 25      | 3        | Would (BC only)                                | 8  |
|                          | Training      | 25      | 3        | Would (ABC only)                               | 8*                                       |
|                          | Training      | 26-27   | 3        | Would (ABC only) Placing 2 coins first         | 33                                       |
|                          | Training      | 28      | 3        | Would (ABC only)                               | 9  |
|                          | 2 Tests       | 29-30   | 4        | Would; Would/Not                               | P;F                                      |
|                          | Training      | 31-32   | 4        | Would/Not                                      | 29                                       |
|                          | 4 Tests       | 33-35   | 5-6      | Would/Not; Text Only; Vertical;<br>Gen. (CD's) | P;P;P<br>P                               |
| Reversal 1               | Training      | 36-37   | 6        | Would  | 83                                       |
|                          | 3 Tests       | 37      | 7        | Would; Would/Not; Gen. (Spoons)                | P;P;P                                    |
| Reversal 2               | Training      | 38      | 7        | Would  | 29                                       |
|                          | 3 Tests       | 38      | 8        | Would; Would/Not; Gen. (Books)                 | P;P;P                                    |

P = Pass; F = Fail; Reading from left to right. \* Indicates that the participant failed to reach the mastery criterion during training.  
Gen.= Generalization test.

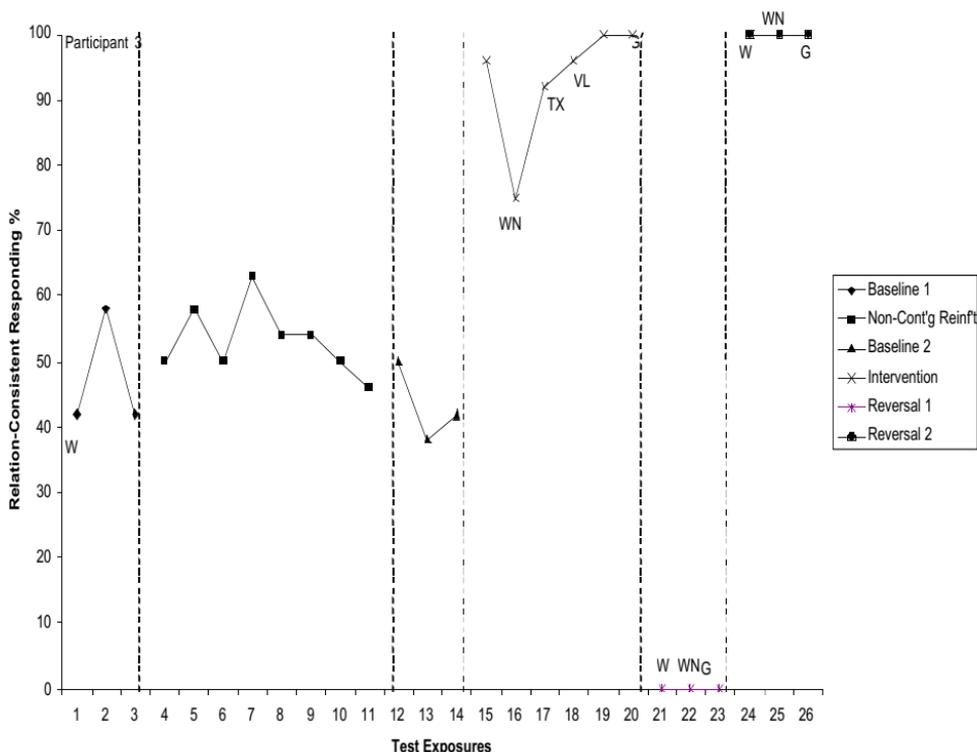


Figure 5. Percentage of relation-consistent responses for Participant 3 on tests of the relations of more-than and less-than.

would and would-not trials (see Figure 4).

### Participant 3

The training and test data for P3 are outlined in Table 3 and the participant's performance during each test exposure is presented in Figure 5. The testing and training procedures employed with this participant were again similar to those employed previously. However, this participant was exposed to a long baseline of non-contingent reinforcement, in order to determine whether extended exposure to the experimental tasks might establish the performance in the absence of contingent reinforcement. A non-contingent reinforcement condition was employed because pilot work had demonstrated that participants invariably found the tasks aversive, and were less willing to cooperate during extended periods conducted in the complete absence of reinforcement.

After failing three baseline tests, P3 was introduced immediately to the extended non-contingent reinforcement training. The number of non-contingent reinforcement trials was based on the number of training and test trials required by P2 to pass the baseline test. Participant 1 required fewer trials than P2, and so using the latter participant

as the basis for determining the minimum number of non-contingent reinforcement trials constituted a stronger test of whether mere exposure to the experimental tasks would generate the performance. Participant 2 required 219 training trials to pass the baseline test. In order to present the training trials in blocks of eight as had been done previously, P3 was exposed to 224 non-contingent reinforcement trials (i.e., 28 blocks of 8). To make this form of training closely resemble the explicit training given to the other participants, similar quantities of reinforcement, trial repetitions, and bead withdrawals to those used previously were employed. For example, in each block of eight trials reinforcement was provided on four trials, two trials were repeated, and a bead was withdrawn after one trial (the sequence of these manipulations was randomized across blocks). The feedback that was provided was entirely random, and may or may not have been correct in terms of the specified relation.

In Sessions 3-7, P3 was exposed to 224 non-contingent reinforcement training trials presented in the manner described above. This training involved only the AB relations because all participants were required to master the AB relations before proceeding to the BC relations. At no point during this training did the participant produce eight consecutively correct responses on the AB relations. On the completion of the 224 trials, he was reexposed to three baseline tests, all of which he failed. Participant 3 was subsequently exposed to the explicit training intervention employed with the previous participants. Similar to P2, he had great difficulty during training of the initial relations. Specifically, he required 121 training trials to reach criterion on the initial AB relations, and 75 training trials to reach criterion on the BC relations. Unlike P2, however, he did eventually reach criterion on both of these types of relations without the use of nonarbitrary interventions. Similar to P1, P3 had particular difficulty during training of the combinatorially entailed ABC relations. After failing to reach criterion on the ABC relations in 40 trials (Sessions 21-23), he was reexposed to training on the AB and BC relations, and reached the mastery criterion quickly. On the second exposure to ABC training, he once again failed to reach the mastery criterion, and an alternative intervention for training these relations was employed.

In Session 26, an intervention in which two coins were placed down first was introduced, and part of the ABC relation was specified. The third coin was then placed down and the complete relation was specified. For example, on one trial, coins A and B were positioned, and the participant was told: "If this coin (A) buys more sweets than this coin (B), which would you choose?" Corrective feedback (but no bead) was provided for a correct response. The third coin was then positioned and the trial was continued with the instructions "If this coin (A) buys more sweets than this coin (B), and this coin (B) buys more sweets than this coin (C), which would you choose?" A bead followed each correct response at this point in the trial and incorrect responses produced no beads. If the child emitted one or two incorrect responses during the trial, the trial-type was repeated, without beads for correct responses. After 33 of these training trials, P3 finally reached criterion on the ABC relations. He was subsequently exposed to training on ABC relations without this intervention and reached criterion in only 9 trials. He then passed an arbitrary baseline test for the first time (Session 29). Similar to P1, he also required explicit training on the would/would-not relations, and subsequently passed

the would/would-not test immediately. He passed all subsequent tests without training. Prior training had taken so long that this participant was not exposed to a follow-up test. He required only minimal training during both contingency reversals (see Figure 5).

#### DISCUSSION

The present study provides evidence that responding in accordance with the relational frame of more-than and less-than is a form of generalized operant behavior. In the current study, all three participants failed to pass baseline tests for responding in accordance with the relations of more-than and less-than. One child (P3) was also provided with an extended baseline of non-contingent reinforcement, but still failed to demonstrate the appropriate relational responding before operant contingencies were introduced. These consistent failures indicated that the target relational performances were not present in the participants' behavioral repertoires. Furthermore, the extensive training required by each of the three children to establish the patterns of relational responding provided even further evidence to support the conclusion that the target relational repertoires were absent prior to the commencement of the study. The data clearly demonstrate that relational responding in accordance with the frame of comparison (i.e., more-than and less-than) was established during the experimental procedures in the behavioral repertoires of the 4 and 5-year old children.

Furthermore, the A-B-A reversal design showed that this form of responding could be brought under operant control. Performance on the generalization tests provided additional evidence of the frame-like qualities of these observed response patterns, in that the children responded relationally to novel stimulus sets (books, CD covers, pencils, spoons, and drinking glasses) and in the context of a novel experimenter. Further evidence of the effectiveness of the current training interventions was demonstrated by the performance of P3, who made no progress during the baseline of non-contingent reinforcement, and yet progressed with the subsequent introduction of the explicit training intervention. Overall, these data provide evidence that responding in accordance with the relational frame of comparison may be established and manipulated as a type of generalized operant behavior.

In the current study, operant contingencies were applied across multiple sets of stimuli and these contingencies successfully established the target relational responses for all three children. Increasingly complex patterns of these relational responses were also established by the operant contingencies (e.g., contextual control by would/would-not), thereby demonstrating that specific patterns of relational responding had been established for each child. Data from the generalization tests, the non-contingent reinforcement condition, and the two contingency reversals also indicated that these relational responses were a class of generalized operant behaviors. The current findings support and extend previous research in this area (Barnes-Holmes, et al. 2001a, 2001b; Healy, et al., 2000).

For each child, the experiment may be considered in terms of two broad stages. The first stage consisted of establishing the basic relational repertoires, whereas the second stage was concerned with increasing the complexity and flexibility of those

repertoires. The results obtained across each of the children during the second stage were relatively consistent. For example, for each child contextual control by would/would-not and control by the two contingency reversals required limited amounts of training. Furthermore, generalization tests across all three children were highly consistent. The first stage of the experiment, however, may appear somewhat more disparate across participants than the latter stage. More specifically, a number of interventions were employed in response to the idiosyncratic relational deficits that emerged for each child. For example, P1 and P3 trained with relative ease on the mutually entailed AB and BC more-than and less-than relations, whereas P2 displayed great difficulty with these relations. The former two children, however, demonstrated some difficulty with the combinatorially entailed relations, whereas P2 did not -once the mutually entailed relations had been established. (Parenthetically, the functional separation of mutual and combinatorial entailment has been reported in a number of previous studies with both children and adults [e.g., Healy, et al., 2000; Lipkens, Hayes, & Hayes, 1993; Pilgrim & Galizio, 1990, 1995; Pilgrim, Chambers, & Galizio, 1995]. The Lipkens, et al. study in particular demonstrated, not unlike the current study, that mutual entailment appears to develop before combinatorial entailment).

In response to these and other individual differences across participants, two key training interventions were employed. For P2, nonarbitrary stimuli (i.e., different numbers of sweets) were used to establish arbitrary relational control. For P3, novel trial-types were employed that integrated both mutual and combinatorially entailed relations within a single trial (i.e., placing the AB coins down first and then immediately after the participant's response, presenting the C coin). Although they may appear disparate, these two interventions fall naturally out of RFT, and indeed are consistent with behavior analytic principles more generally (see final paragraph).

In the current study, the operant contingencies were designed to establish contextual functions that we would expect to be acquired, sooner or later, through each child's normal interactions with the English-speaking verbal community. In other words, we assume that the Crel functions that were established in this study for the terms "more-than," and "less-than," would have been acquired eventually during the course of each child's normal development. The use of "real words" in this way could be criticized on the grounds that natural learning (in the extra-experimental environment) may have in some undefined way facilitated the performances obtained during the study. Although this remains a possibility, it seems unlikely that natural learning played a significant role in generating the very specific and complex performances observed in the current study. Furthermore, the fact that one participant was provided with an extended baseline of non-contingent reinforcement, and still required extensive training thereafter, seriously undermines the plausibility of a natural learning explanation for the current data. But why did we use real words in the current study? Relational Frame Theory constitutes a modern behavioral approach to human language and cognition (Hayes, et al., 2001), and thus it was deemed important for research in this area to begin to work directly with natural language itself. Of course, whenever laboratory research aims to make direct contact with the natural environment, experimental precision is very often traded for ecological validity. At this point in the RFT research program, the current shift

towards greater ecological validity seemed important.

Perhaps the most critical feature of the current data is the insight they provide into the extent and nature of the training history that is required to establish responding in accordance with relational frames. Some researchers have asked, for example, whether explicit training in mutual and combinatorial entailment is required in order to establish responding in accordance with these two properties, or whether training in mutual entailment alone will suffice (e.g., Boelens, 1994; Horne & Lowe, 1996). The RFT perspective on this question was nicely summarized by Hayes and Wilson (1996):

How much and what kind of training is needed for generalization of a relational response is an empirical matter. However, the general logic of RFT suggests that at least some direct training in combining relations (e.g., both  $A \rightarrow C$  and  $C \rightarrow A$  training [following  $A \rightarrow B$ ,  $B \rightarrow C$ ,  $B \rightarrow A$ , and  $C \rightarrow B$  training]) is necessary. Using RFT terms, this point has been made explicitly in early expositions; for example equivalence emerges because “mutual entailment, combinatorial entailment and transfer of functions are *directly* trained” (Hayes, 1991, p. 25). It is important to note here that combinatorial entailment subsumes both  $A \rightarrow C$  and  $C \rightarrow A$  relations. . . It does seem likely, however, that once the most basic relational unit is established through training in mutual and combinatorial entailment, relatively fewer trained instances of combinatorial entailment will be needed to build out this relational response. Were it not true, every level of relational complexity (e.g., with larger and larger sets of related stimuli) might have to be arduously trained. Consider, for instance, a case in which one was taught to select B in the presence of A, C in the presence of B, D in the presence of C, and so on to the 100th node. We doubt that an individual would have to have a history of direct training to match the 100th stimulus to the 1st, the 100th to the 2nd, the 98th to the 1st... and so on for all possible transitive and equivalence relations among the 100 stimuli. *At some point, RFT would predict that the operant of combining relations would itself generalize* (emphasis added, p. 227).

Clearly, the current data provide support for the foregoing interpretation. Indeed, other evidence obtained from a related study in which several of the current authors established complex patterns of responding in accordance with the relational frame of opposite in children of a similar age provided further evidence for this interpretation (Barnes-Holmes, Barnes-Holmes, & Smeets, this volume). For example, one participant in that study required explicit training in the relation of opposite using two, three, four, and five coins before responding in accordance with opposite generalized, without explicit training, to six, seven, and eight coins/objects. Similar effects were also observed with the other two participants who participated in that study.

A possibly important issue arising from the current research is that the participants were already demonstrating a relatively advanced level of language ability before entering the study, and this may have played a critical role in generating the observed relational performances. Certainly, the relative ease with which ‘would-not’ control was established for all of the participants, suggests that preexperimental verbal skills were indeed important. One theoretical or interpretive problem that arises at this point is the possibility that very different behavioral processes were involved in the establishment of the language

skills with which the children entered the study, than the operant processes that were the focus of the current research. Insofar as this was the case, this would limit the theoretical implications of the current work, vis-a-vis RFT's analysis of human language and cognition. At the present time, however, there appears to be no reason to suspect that fundamentally different behavioral processes were involved in the participants' preexperimental language learning. In effect, the conservative and parsimonious assumption at this point is that the same operant processes that were used in the current study to establish the specific verbal or relational skills were also heavily involved in the establishment of the language skills that the children possessed before entering the study. Of course, further research will be needed to determine whether this conservative assumption proves to be correct. In any case, from a purely applied perspective, the current data clearly indicate the possible utility in adopting an RFT, operant approach to the establishment of generalized verbal or cognitive skills in young children.

The present study was clearly generated by RFT, but alternative interpretations of the current data are possible. For example, it could be argued that the children learned initially to respond to the first coin when given the spoken word "More," and to respond to the last coin when given "Less." In other words, the training procedures established two stimulus classes, with the S+ and S- functions of the two classes determined by the words "More" and "Less." One might argue that a functionally similar effect was reported by Vaughan (1988) who established two stimulus classes by means of repeated reversal training with pigeons. Nevertheless, it is important to recognize that Vaughan employed the same stimuli throughout the entire study, and thus the two stimulus classes were directly trained (see Hayes, 1989). In the current study, however, many novel stimulus sets were introduced and responding to these sets came under the contextual control of "More" and "Less" in the absence of explicit reinforcement. Furthermore, this performance itself came under the contextual control of "Would" and "Would-Not," and finally these contextually controlled performances were successfully manipulated across two contingency reversals. Even if one chooses not to interpret these data in terms of RFT, the results of this study do significantly extend the findings reported by Vaughan (1988).

Notwithstanding the foregoing argument, RFT points to complexities that are not immediately apparent using class-based interpretations of the current data. For example, the relational performances obtained in the current study constitute only a small number of the possible response patterns that define a relational frame. Consider, for instance, the frame of more-than and less-than. As indicated in the Introduction, with three elements a total of 12 different trial-types could be constructed using our linear problem-solving task. Furthermore, this number increases significantly if both linear and nonlinear sequences with the A, B, and C coins/objects are presented in all possible positions (e.g.,  $B > A < C$ , etc.). If even one or two additional elements are then added, this number increases dramatically (Newstead, Evans, & Byrne, 1993). Clearly, an exhaustive analysis of the more-than/less-than relational frame is far from a simple matter, but it seems unlikely that an individual would have to be trained on all possible trial-types in order to produce all possible response patterns in accordance with a particular relational frame. What subset of trial-types must be taught is an empirical matter and has indeed

been the source of a recent program of research undertaken by some of the current authors (see Vitale, Barnes-Holmes, & Barnes-Holmes, 2003).

The foregoing points to one possible criticism of the present study. Because the children were trained and tested on the same four trial-types using only two or three coins, one might argue that the test performances, even on novel stimulus sets, were not indicative of *derived* relational responding (i.e., the children were exposed to novel stimuli, but not novel trial-types, during the tests). In transitive inference research (Bryant & Trabasso, 1971; Russell, McCormack, Robinson, & Lillis, 1996; von Fersen, Wynne, Delius, & Staddon, 1991; see also Green, et al., 1993, for an excellent review), for example, participants may be trained on a number of stimulus pairs (e.g.,  $A > B$ ;  $B > C$ ;  $C > D$ ;  $D > E$ ) and then tested on two or more nonadjacent pairs ( $B > D$ ?). In this case, different trial-types are used across training and testing, and thus one might argue that the test performances are genuinely derived. In any case, it is important to recognize that the current study was primarily concerned with developing procedures for establishing and manipulating relational performances, as defined by RFT, and demonstrating their generalized operant-like qualities, rather than testing a particular cognitive or nonhuman model of transitive inference.

At a more general level, approaching relational responding as generalized operant behavior may provide new and possibly useful ways of conceptualizing human language and cognition (Hayes, et al., 2001). From the perspective of RFT, relational activities are considered to be the functional-analytic bedrock of human cognitive and verbal abilities. This behavior analytic view avoids the typical approach taken by cognitive psychology, which has tended to emphasize “content” by focusing on specific words and/or the acquisition of specific concepts applicable in the real world. For RFT, the key focus should be on the relational activities *per se*, rather than on particular words or concepts. In the present study, for example, large numbers of “pretend” coins, and a range of randomly selected objects, were used to establish generic patterns of relational responding. Perhaps a similar approach could be taken in educational settings in which learners are trained in both real world concepts and in various types of relational responding. 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 opposite of Y, and Y is the same as Z, do I like Z if I like X?” Of course, broadly similar training does occur during the course of normal educational practice. However, such practice is not designed specifically to target the key cognitive or relational skills (see Fredrick, Deitz, Bryceland, & Hummel, 2000). In contrast, RFT is directly concerned with these core relational skills and how they might be harnessed for bringing about improved educational achievement. What is most exciting about this research agenda is that the same general process of relational framing may be at the heart of a very wide range of cognitive abilities.

#### REFERENCES

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P.M. (2001a). Exemplar training and a

derived transformation of function in accordance with symmetry. *The Psychological Record*, 51, 287-308.

- Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P.M. (2001b). Exemplar training and a derived transformation of function in accordance with symmetry: II. *The Psychological Record*, 51, 589-604.
- Barnes-Holmes, D., Dymond, S., Roche, B., & Gray, I. (1999). Language and cognition. *The Psychologist*, 12, 500-504.
- Barnes-Holmes, Y., Barnes-Holmes, D., & Smeets, P.M. (2004). Establishing relational responding in accordance with opposite as generalized operant behavior in young children. *International Journal of Psychology and Psychological Therapy*, 4, 559-585.
- Boelens, H. (1994). A traditional account of stimulus equivalence. *The Psychological Record*, 44, 587-605.
- Bryant, P.E. & Trabasso, T. (1971). Transitive inference and memory in young children. *Nature*, 232, 456-485.
- Devany, J.M., Hayes, S.C., & Nelson, R.O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior*, 46, 243-257.
- Dymond, S. & Barnes, D. (1995). A transformation of self-discrimination response 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.
- 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.
- von Fersen, L., Wynne, C.D.D., Delius, J.D., & Staddon, J.E.R. (1991). Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behaviour Processes*, 17, 334-341.
- Fredrick, L.D., Deitz, S.M., Bryceland, J.A., & Hummel, J.H. (2000). *Behavior analysis, education, and schooling*. Reno, NV: Context Press.
- Green, G., Stromer, R., & Mackay, H.A. (1993). Relational learning in stimulus sequences. *The Psychological Record*, 43, 599-616.
- Greer, R.D., Phelan, C.S., & Sales, C. (1993). *A costs-benefits analysis of a graduate course*. Paper presented at the International Conference of the Association for Behavior Analysis, Chicago, IL, USA.
- Hayes, S.C. (1989). *Rule-governed behavior: Cognition, contingencies, and instructional control*. New York: Plenum.
- Hayes, S.C. (1991). A relational control theory of stimulus equivalence. In L.J. Hayes & P.N. Chase (Eds.), *Dialogues on verbal behavior: The first international institute on verbal relations*. Reno: Context Press.
- Hayes, S.C. (1996). Making sense of spirituality. *Behaviorism*, 12, 99-110.
- Hayes, S.C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York: Plenum.
- Hayes, S.C. & Wilson, K.G. (1996). Criticisms of relational frame theory: Implications for a behavior analytic account of derived stimulus relations. *The Psychological Record*, 46, 221-236.
- Healy, O., Barnes-Holmes, D., & Smeets, P.M. (2000). Derived relational responding as generalized operant behavior. *Journal of the Experimental Analysis of Behavior*, 74, 207-227.
- Horne, P.J. & Lowe, C.F. (1996). On the origins of naming and other symbolic behavior. *Journal of the*

*Experimental Analysis of Behavior*, 65, 185-241.

- Lipkens, R., Hayes, S.C., & Hayes, L.J. (1993). Longitudinal study of the development of derived relations in an infant. *Journal of Experimental Child Psychology*, 56, 201-239.
- Newstead, S.E., Evans, J.St.B.T., & Byrne, R.M.J. (1993). Relational inferences. In S.E. Newstead, J.St.B.T. Evans, & R.M.J. Byrne (Eds.). *Human reasoning: The psychology of deduction*. London: Psychology Press.
- Pilgrim, C., Chambers, L., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: II. Children. *Journal of the Experimental Analysis of Behavior*, 63, 239-254.
- Pilgrim, C. & Galizio, M. (1990). Relations between baseline contingencies and equivalence probe performances. *Journal of the Experimental Analysis of Behavior*, 54, 213-224.
- Pilgrim, C. & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: I. Adults. *Journal of the Experimental Analysis of Behavior*, 63, 225-238.
- Rehfeldt, R.A. & Hayes, L.J. (2000). The long-term retention of generalized equivalence classes. *The Psychological Record*, 50, 405-428.
- Roche, B. & Barnes, D. (1996). Arbitrarily applicable relational responding and human sexual categorization: A critical test of the derived difference relation. *The Psychological Record*, 46, 451-475.
- Roche, B. & Barnes, D. (1997). A transformation of respondently conditioned stimulus function in accordance with arbitrarily applicable relations. *Journal of the Experimental Analysis of Behavior*, 67, 275-300.
- 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.
- Russell, J., McCormack, T., Robinson, J., & Lillis, G. (1996). Logical (versus associative) performance on transitive reasoning tasks by children: Implications for the status of animals' performance. *The Quarterly Journal of Experimental Psychology*, 49b, 231-244.
- Saunders, R.R., Wachter, J., & Spradlin, J.E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discrimination procedures. *Journal of the Experimental Analysis of Behavior*, 49, 95-115.
- Steele, D. & Hayes, S.C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior*, 56, 519-555.
- Vaughan, W. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 36-42.
- Vitale, A., Barnes-Holmes, Y., & Barnes-Holmes, D. (2003). *Facilitating complex patterns of responding in accordance with the frame of comparison in adults*. Paper presented at the annual convention of The Association for Behavior Analysis, San Francisco, USA.
- Wulfert, E. & Hayes, S.C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior*, 50, 125-144.

Received October 15, 2004

Final acceptance October 30, 2004