# Facilitating Responding in Accordance with the Relational Frame of Comparison: Systematic Empirical Analyses

**4 authors**, including:

Yvonne Barnes-Holmes
Ghent University
**169** PUBLICATIONS **6,424** CITATIONS

SEE PROFILE

Claire Campbell
Ulster University
**21** PUBLICATIONS **208** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project    See here for some new projects View project

Project    A Study of Brexit and Social Identity in Northern Ireland View project

# FACILITATING RESPONDING IN ACCORDANCE WITH THE RELATIONAL FRAME OF COMPARISON: SYSTEMATIC EMPIRICAL ANALYSES

Agata Vitale

*Mary Immaculate College, Limerick*

Yvonne Barnes-Holmes, Dermot Barnes-Holmes, and Claire Campbell

*National University of Ireland, Maynooth*

*The current article examines patterns of adult responding to different types of more-than and less-than relations, as well as procedures for facilitating responding in accordance with these relations. Using parameters suggested in the three-term series literature, the more-than and less-than relations were separated into six distinct trial types. Systematic comparisons were then drawn between repeated test exposures, the provision of automated feedback, and the presentation of nonarbitrary trials. The results showed that feedback was the most effective intervention for facilitating responding to the target relations. When feedback and nonarbitrary trials were combined, their impact was only marginally better than either intervention alone. These findings are discussed in terms of relational frame theory and interventions for remediating deficits in derived relational responding.*

The concept of emergent stimulus relations has been of interest to behavior analysts for over 30 years (e.g., Sidman, 1971). This interest has led to the development of a number of theoretical approaches that offer different definitions of stimulus relations and different accounts of how they emerge naturally or how they can be established for purposes of experimentation or interventions. For example, Sidman (1994) offered a mathematical set theory of stimulus equivalence in terms of reflexivity, symmetry, and transitivity. In contrast, with relational frame theory (RFT), the multiple types of stimulus relations (beyond equivalence per se) were of more concern, and the role of multiple exemplar training has been emphasized in their development or establishment (Hayes, Barnes-Holmes, & Roche, 2001).

According to RFT, sophisticated verbal behavior is composed of the ability to respond in accordance with a range of derived stimulus relations, including coordination, difference, comparison, hierarchy, and even deictic frames of perspective taking (Barnes-Holmes, Hayes, Dymond, & O'Hora, 2001). Numerous studies to date have provided strong empirical support for the existence of relational repertoires as suggested by RFT, and for their operant nature (e.g., Dymond & Barnes, 1995, 1996; Steele & Hayes, 1991; Roche & Barnes, 1996, 1997; Wulfert & Hayes, 1988). However, demonstration research alone does not identify the mechanisms by which relational frames develop or can be enhanced if they are found to be deficient.

A number of studies have successfully demonstrated the natural development and establishment or facilitation of relational repertoires both with children and with adults (e.g., Lipkens, Hayes, & Hayes, 1993). In one study with young normally developing children, for example, exemplar training proved to be more effective than name training in facilitating transformations of function in accordance with symmetry (Barnes-Holmes, Barnes-Holmes, Roche, & Smeets, 2001a; see also Barnes-Holmes, Barnes-Holmes, Roche, & Smeets, 2001b; Horne & Lowe, 1996; Horne & Lowe, 2007). However, for RFT, different interventions may be needed to facilitate different types of relational responding. For example, Barnes-Holmes, Barnes-Holmes, Smeets, Strand, and Friman (2004) integrated nonarbitrary and arbitrary relations to establish arbitrary comparative relations with young children. In another study with adults, researchers reported positive benefits for written feedback in facilitating coordination relations presented across novel stimulus sets (Healy, Barnes-Holmes, & Smeets, 2000).

As a behavioral theory, RFT would predict the success of a range of standard behavioral interventions, as well as those driven more directly by the theory itself, in the establishment or facilitation of relational repertoires. In the former case, for example, the theory would predict that feedback across stimulus sets would be successful if it functioned appropriately as a type of exemplar training of all aspects of the target relational response. In the latter case, for example, RFT would predict that the absence of nonarbitrary relations would make it almost impossible to establish the arbitrary counterparts of those relations. In any case, these remain empirical issues.

One area within the field of psychology that may provide answers to these questions is work by cognitive psychologists on *the three-term series problem,* which appears to overlap considerably with the RFT description of the relational frame of comparison. The three-term series problem consists of three related elements (e.g., A, B, and C) described within two premises (e.g., A is related to B; B is related to C), followed by a question concerning the relationship between two of the elements (e.g., A is related to B; B is related to C. How is A related to C?).

Cognitive researchers of three-term series problems have devoted considerable attention to task-based variables that render problems more or less difficult to solve, including the organization of the premises and the use of specific relational terms. For example, some tasks present the *same* comparative terms in both premises (e.g., Ken has *more than* Bill; Bill has *more than* Dave—denoted as A > B > C), whereas others employ mixed comparatives (e.g., Ken has *more than* Bill; Bill has *less than* Dave—A > B; B < C). According to Hunter (1957), the former problems are easier to solve than the latter because the two comparatives employed are the same. Hence, the A > B; B < C problem might be made easier if the comparative in the second premise is reversed

from the B < (less-than) C relation to a C > (more-than) B relation so that the two premises now contained a "more-than" relation.

Another feature of three-term series problems that affects task difficulty is the degree of specificity of the information in the two premises. That is, some three-term series problems have a clearly identifiable solution (these are referred to as *specified* problems—Sternberg, 1981), whereas others do not (these are referred to as *unspecified* problems—Hayes-Roth & Hayes-Roth, 1975). With unspecified problems, the information contained within the premises is not adequate so that not all the relations can be identified. For example, consider the problem A > B, B > C, in which all relations among the three terms are specified (i.e., one can derive that B < A; C < B; A > C; and C < A). Now consider the problem A > C; B > C. This problem is unspecified because a number of conflicting inferences can be generated from the information provided. Specifically, A may be more than B or B may be more than A; therefore, the A-B relation remains unspecified. In this case, a correct answer requires discriminating that the problem cannot be solved. Although some cognitive researchers have reported more errors on unspecified than specified problems (Clark, 1969; Hayes-Roth & Hayes-Roth, 1975; Moeser & Tarrant, 1977; see also Byrne & Johnson-Laird, 1989), others have found the converse (e.g., Sternberg, 1981).

One explanation for these conflicting findings was suggested by Clark (1969), who analyzed combinations of same and mixed comparatives within specified and unspecified problems. He concluded that the combination of unspecified and mixed relations generates problems that are particularly difficult. As a result, unspecified tasks that involve same relations may be easier to solve than specified tasks involving mixed relations (see Mani & Johnson-Laird, 1982). However, consistent with the generic aims of cognitive psychology and a strong emphasis on theory building over manipulation or intervention, the cognitive literature contains almost no empirical studies that have attempted to remediate weaknesses in the solution of three-term series problems, and no predictions about what may or may not be effective there.

Investigations into transitive relations have also formed a core interest of cognitive researchers of three-term series problems. According to Hunter (1957), nontransitive relations contain two premises with the repeated term in the middle (e.g., A < B, B < C), and these premises explicitly state relationships involving the repeated term (i.e., all those involving B). As a result, nontransitive relations do not state the relationship between nonadjacent and nonrepeated terms (e.g., between A and C). Although a number of definitions of transitive relations are found within the cognitive literature, one key feature is the requirement to state relations between nonadjacent terms. This is the definition adopted in the current article. Hence, in the current research the problem "A > B; B > C" is referred to as nontransitive, because it does not state a relationship between A and C. In contrast, the problem "A > B; C > A" does contain a transitive premise (C > A) with the relationship between nonadjacent terms stated explicitly.

For Hunter, nontransitive relations are easier to solve than transitive relations, because the format of the premises in nontransitive relations is linear (e.g., A < B, B < C) and therefore can be worked through in a single sequence. Transitive relations (e.g., A < B, C < B), on the other hand, are nonlinear and must be converted to linear relations before they can be solved. As a result, they are more difficult. Numerous researchers (Russell,

McCormack, Robinson, & Lillis, 1996) have identified linearity as a key feature in the solution of transitive inference problems. According to Hunter, linear problems are easier to solve because they reflect the "natural cognitive order" in which humans readily deal with the information provided.

The current research, although designed from concepts and methodologies associated with behavioral psychology, overlaps significantly with work on the three-term series problem generated within the mainstream cognitive literature. The current research drew heavily on the cognitive tradition in attempting to generate relational problem-solving tasks that normally developing adults would find difficult and which would therefore generate weak relational performances. For example, the current research involved the development of three-term series problems involving same, mixed, specified, unspecified, and transitive relations. Interventions were then drawn from the traditional behavioral literature, as well as from research on RFT, for problems in which weak relational performances were observed. For example, the work investigated the utility of repeated exposure, written feedback, and nonarbitrary trials as interventions. Specifically, in Experiment 1 a repeated exposure methodology was examined to test the assumption that simply increasing the number of exposures to the trial types might enhance the target relational performances. In Experiment 2, feedback was used as a training methodology in a limited number of trials. In Experiment 3, nonarbitrary trial types were employed in a restricted number of trials on the basis of findings from previous research with children (see Barnes-Homes et al., 2004). When this methodology generated only limited improvements, the amount of exposure to nonarbitrary trials was increased because of the possibility that limiting the training to only Phase 2 was inadequate for generating competent arbitrary performances. In Experiment 4, therefore, participants received repeated exposures to nonarbitrary trial types. In Experiment 5, the previous interventions of feedback and nonarbitrary trials were combined in only Phase 2, to determine whether integrating the two interventions would generate a stronger intervention regime than either intervention alone.

## Experiment 1

### *Method*

*Participants.* A total of 10 naïve undergraduate students of the National University of Ireland, Maynooth (NUIM), participated in Experiment 1. All were between 18 and 24 years old and were recruited through class announcements and notice boards. None of the participants received any remuneration for their involvement in the experiment, and all participated individually.

*Setting.* Experiment 1 was conducted in an experimental room in the Department of Psychology at NUIM. The room was in a quiet location and generally free from distraction. During the automated procedure, the participant remained alone in the experimental room while the experimenter was seated outside. Each participant was exposed to three experimental phases. Only one phase was presented per day, with the two subsequent phases presented on subsequent consecutive days, availability permitting. Each experimental phase lasted approximately 65 min, so that participation in the full experiment lasted approximately 195 min.

*Apparatus.* The apparatus used in Experiment 1 involved an automated procedure that ran on an Apple iBook laptop computer with a Power PC G3 500 MHz processor, a 12.1-in. LCD screen, and a standard computer mouse. All stimulus presentations and participant responses within the automated procedure were recorded by a Visual Basic Program (Version 6) presented through Real PC (Version '95). Each automated trial depicted three identically sized circles (referred to as "coins") that differed only in color—one red, one blue, and one yellow. For experimental purposes, the three coins were designated as A (red), B (blue), and C (yellow), although participants never saw these labels. Each trial also depicted three brown rectangles (referred to as "coffee containers," which differed only in size—one large (referred to as "full of coffee"), one medium ("half full of coffee"), and one small ("a little coffee"). During each trial, a small black tin was also present on the screen. This item was labeled the "I cannot know" tin.

*Trial types.* The automated procedure contained a total of 48 trials, divided according to six trial types that differed according to three basic dimensions: (1) the *target relations* stated among the coins; (2) whether the relations between the coins were to be derived as *specified* or *unspecified*; and (3) whether or not the relations to be derived between the coins were *transitive*. Table 1 presents the full list of trial types and their relevant categorizations.

Table 1
*The Six Trial Types Presented in Experiment 1*

| Specified-Same Relations | | Unspecified-Same Relations | |
|---|---|---|---|
| MORE-MORE | LESS-LESS | MORE-MORE | LESS-LESS |
| A > B; B > C | A < B; B < C | A > B; C > B | A < B; C < B |
| B > C; A > B | B < C; A < B | B > A; B > C | B < A; B < C |
| B > A; C > B | B < A; C < B | B > C; B > A | B < C; B < A |
| C > B; B > A | C < B; B < A | C > B; A > B | C < B; A < B |

| Specified-Mixed Relations | | Unspecified-Mixed Relations | |
|---|---|---|---|
| MORE-LESS | LESS-MORE | MORE-LESS | LESS-MORE |
| A > B; C < B | A < B; C > B | A > B; B < C | A < B; B > C |
| B > C; B < A | B < C; B > A | B > C; A < B | B < C; A > B |
| B > A; B < C | B < A; B > C | B > A; C < B | B < A; C > B |
| C > B; A < B | C < B; A > B | C > B; B < A | C < B; B > A |

| Specified-Same Transitive Relations | | Unspecified-Mixed Transitive Relations | |
|---|---|---|---|
| MORE-MORE | LESS-LESS | MORE-LESS | LESS-MORE |
| A > B; C > A | A < B; C < A | A > B; C < A | A < B; C > A |
| A > C; B > A | C < A; B < A | A > C; B < A | A < C; B > A |
| C > B; A > C | C < B; A < C | C > B; A < C | C < B; A > C |
| C > A; B > C | C < A; B < C | C > A; B < C | C < A; B > C |

Each trial constituted a three-term series problem that required participants to try to determine the relations between all three coins. During some trial types, the relations presented across the two premises were the same. For example, participants may have been presented with a task that contained two more-than relations: "Coin A is worth more than coin B, and coin B is worth more than coin C" (denoted as A > B; B > C). Alternatively, two less-than relations may have been presented (e.g., B < C; A < B). In contrast, some trials presented mixed relations across the premises, in which one premise contained a more-than relation, whereas the other contained a less-than relation or vice versa (e.g., B > C; A < B or C < B; B > A). Trials that contained two more-than or two less-than relations were referred to as *same* trials, whereas trials that contained one more-than and one less-than relation were referred to as *mixed* trials.

During some trial types, the information presented across the two premises was sufficient to allow participants to correctly determine the remaining unstated relations between the coins. For example, if presented with A > B; B > C, one could correctly derive the relations A > C and C < A without additional information. In this case, the trials were referred to as *specified*. In contrast, during *unspecified* trials the information presented was not sufficient to allow the participant to derive the target relations. For example, if presented with A > B; B < C, one cannot determine the relations between A and C. Each test exposure contained both specified-same and specified-mixed trial types.

The trials were also differentiated in terms of whether or not a relation between nonadjacent coins (i.e., a transitive relation) was presented. For example, in the trial C < A; B < A, the first premise identifies a transitive relation between the nonadjacent coins A and C. In contrast, the trial B > C; A < B presents relations between only adjacent coins B and C as well as A and B.

In summary, the presentation of same, mixed, specified, unspecified, and transitive relations generated a total of six basic trial types that were presented within each test. Specifically, each test contained nontransitive specified-same, unspecified-same, specified-mixed, and unspecified-mixed trial types, as well as two transitive trial types—specified-same transitive and unspecified-mixed transitive (see Table 1).

*Procedure.* Participants were exposed to the automated procedure twice during each of the three experimental phases (i.e., they each received a total of six exposures), with a 5-min break between each exposure. At the beginning of the automated procedure, each participant was provided with a printed set of general instructions as follows:

> During the experiment, the computer will present you with a number of problems to solve. For each problem, three coins, three coffee jars (with coffee), and a black container will be presented. Each coin is worth the amount of coffee in one of the three containers. Each time the computer will tell you about the relative value of each of the three coins. Your task is to work out which coin is worth which jar of coffee and then drag and drop each coin into the space below the appropriate jar. For some questions it is impossible to know where two of the three coins should be placed. When this is the case, drop

these two coins into the space below the black container and drop the remaining coin into the space below the appropriate coffee jar.

The experimenter left the room shortly after the instructional phase and the automated procedure commenced immediately. Participants simply clicked on an intermediate screen to indicate their readiness to proceed.

The automated procedure contained 48 test screens, each presenting a single trial with the three coins (red, blue, and yellow), the three coffee jars (small, medium, and large), and a black tin (the "I cannot know" tin). On all trials, the three coins were presented in fixed locations (A/red left, B/blue middle, and C/yellow right), as were the three jars (small/left, medium/middle, and large/right). The basic task required participants to drag and drop each coin under the appropriately sized coffee jar or into the black tin depending upon the relations between the coins as stated during the trial.

For illustrative purposes, consider the specified-same trial denoted as A < B; B < C. During this trial, the following instruction appeared on screen: "The red coin (A) is worth less than the blue coin (B); and the blue coin (B) is worth less than the yellow coin (C)" (again participants did not see the alphanumeric labels). In this case, a correct response was the participant's deriving the following information and placing the coins accordingly. The red coin (A) is worth the least (because A is worth less than B and B is worth less than C) and so it should be placed with the smallest jar. Coin B is in the middle (because B is worth more than A but less than C) and so the blue coin should be placed with the medium-sized jar. Coin C is the largest (because C is worth more than both A and B) and so the yellow coin should be placed with the largest jar. Given that the above example depicted a specified trial, no coins would be dragged to the black tin, because the three coins could readily be placed beside the three jars.

Now consider the unspecified-mixed trial "C > B; B < A." During this trial, the following instruction appeared on screen: "The yellow coin (C) is worth more than the blue coin (B); and the blue coin (B) is worth less than the red coin (A)." In this case, a correct response was the participant's deriving the following information and placing the coins accordingly. The blue coin (B) is worth less than the red coin but also worth less than the yellow coin, so B goes beside the smallest jar. However, in this case one cannot determine the precise relationship between A and C so that the red and yellow coins must be placed in the black tin (because one cannot determine which jars they should go to).

A number of additional minor instructions were presented on each trial to allow participants to interact with the program freely and correctly, and to proceed appropriately through the experiment. For example, an additional button box was located at the bottom of the screen, with the words "Start Again," which returned all of the coins to their original locations and so allowed the participant to begin the trial again. Similarly, a second "Finish Trial" button box enabled the participant to proceed immediately to the next screen only after all three coins had been moved either to a jar or to the black tin. On each trial, participants could place

only one coin in each jar but could place any number of coins (including none) in the black tin (the program did not permit alternative responses ). No feedback was presented after any trial.

At the end of the final test trial, the program thanked participants for their time so far and advised them to contact the experimenter seated outside the room. They were then given a 5-min break and were reexposed to the entire test for a second time. These two consecutive test exposures constituted Phase 1 of the experiment. Phases 2 and 3 were identical to Phase 1, so that each participant was exposed to a total of six test exposures.

## Results and Discussion

Although all participants completed 2 test exposures in each phase, the data from only the 3- second exposures were analyzed because pilot work had indicated that responding during the first exposure was often erratic. The accuracies of participants' performances were grouped according to the six relation types, and these are presented in Figure 1.
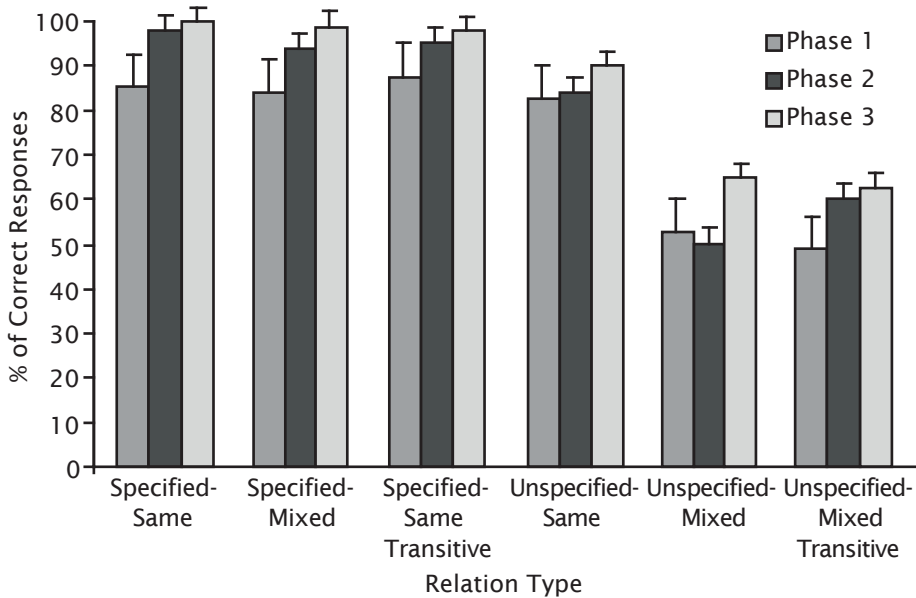


Figure 1. The percentage of correct responses on the six relation types presented in the second test exposure of Phases 1, 2, and 3 of Experiment 1.

Participants overall performed better on the specified than on the unspecified relations, with the weakest performances recorded on the unspecified-mixed and unspecified-mixed transitive relations. Although initially strong, performances on the three specified relations (specified-same, specified-mixed, and specified-mixed transitive) improved further across phases, with a mean accuracy overall of 98.8% across the three types by Phase 3. However, on the weakest unspecified-mixed and unspecified-mixed transitive relations, the mean accuracy across phases

changed little and was not high at any point. Specifically, accuracy on these relations ranged from 52.7% and 48.9% in Phase 1; 50.2% and 60.2% in Phase 2; and 65.1% and 62.7% in Phase 3, respectively. Incidentally, performances on unspecified-same relations were generally more similar to specified than unspecified performances. That is, the mean accuracies recorded for these relations were as follows: 82.6% in Phase 1; 83.9% in Phase 2; and 90% in Phase 3.

A $6 \times 3$ repeated measures Analysis of Variance (ANOVA) was conducted with both relation type and phase as within participant variables. This analysis revealed significant effects for both variables: relation type, $F(5,45) = 14.188$, $p < .0001$, $\eta_p^2 = 0.612$, and phase $F(2,18) = 4.556$, $p < .0251$, $\eta_p^2 = 0.336$, but no interaction effect, $F(10,90) = 0.524$, $p = .8687$, $\eta_p^2 = 0.055$. Post hoc analyses (Fisher's PLSD) revealed significant superiorities of specified over unspecified responding across all phases, particularly when the unspecified responding contained mixed and/or transitive relations. Table 2 summarizes these differences. The consistency of responding on the various trial types across phases indicated that participants' performances on specified relations remained highly accurate, whereas responding to unspecified-mixed and unspecified-mixed transitive relations remained weak.

Table 2
*Statistical Comparisons of Participant's Performances on Each Relation Type Across All Three Phases Presented in Experiment 1*

| Relation Type Comparisons | Phase 1 | Phase 2 | Phase 3 |
|---|---|---|---|
| Specified-same vs. Unspecified-mixed | .0007 | <.0001 | <.0001 |
| Specified-same vs. Unspecified-mixed transitive | .0002 | .0001 | <.0001 |
| Specified-mixed vs. Unspecified-mixed | .0011 | <.0001 | .0001 |
| Specified-mixed vs. Unspecified-mixed transitive | .0003 | .0005 | <.0001 |
| Specified-mixed transitive vs. Unspecified-mixed | .0003 | <.0001 | .0002 |
| Specified-mixed transitive vs. Unspecified-mixed transitive | <.0001 | .0003 | <.0001 |
| Unspecified-same vs. Unspecified-mixed | .0017 | .0005 | .0036 |
| Unspecified-same vs. Unspecified-mixed transitive | .0005 | .0116 | .0015 |

Participants in Experiment 1 received a total of 6 complete exposures to the test protocol and from the outset produced significantly better performances on specified versus unspecified relations, particularly when the unspecified relations were mixed and/or transitive. Although the strong performances on specified relations improved further, the unspecified-mixed and unspecified-mixed transitive relations remained significantly weaker. This finding suggested that although some improvement resulted from repeated exposure alone, greater improvement—especially in the weaker relations— would appear to require more explicit forms of intervention. This issue was addressed in Experiment 2.

## Experiment 2

The limited impact of repeated test exposure on the target relational performances in Experiment 1 prompted the search for more explicit forms of intervention that might be more effective here. Previous research by Healy et al. (2004) reported the utility of written feedback in the facilitation of coordination relations across multiple stimulus sets with adult participants. In line with this evidence, Experiment 2 incorporated written feedback into the program employed in Experiment 1.

### Method

*Participants.* Experiment 2 involved 10 naïve participants who were undergraduate students at NUIM. Participants were 18–24 years old, and they were recruited through class announcements and university notice boards. No participants received any remuneration for their involvement in the experiment.

*Apparatus.* The automated program employed in Experiment 1 was modified for its use in Experiment 2 with the inclusion of automated feedback after every trial in only Phase 2 (for both exposures).

*Procedure.* The sequence of Experiment 2 was identical to that in Experiment 1 except that an ABA design was employed. That is, Phase 1 was composed of 2 test exposures, Phase 2 was composed of 2 training exposures (i.e., test exposures plus feedback), and Phase 3 again was composed of 2 test exposures. This design was employed to determine the specific impact of the feedback intervention on the subsequent test performances relative to the test performances prior to feedback. All other aspects of the procedure in Experiment 2, including the instructions given, were identical to those in Experiment 1.

In Experiment 2, feedback was provided after all trials in Phase 2. That is, if a participant responded correctly, the word "Correct" appeared immediately at the bottom of the screen. If a participant responded incorrectly, the word "Wrong" similarly appeared. Both types of feedback remained on screen for 7 s. During this time, the "Start Again" and "Finish Trial" buttons were inactive so that participants had adequate exposure to the feedback, and they could not proceed to the next trial during the presentation of feedback. At the end of the 7 s, the intermediary screen appeared, and participants could proceed to the next trial. Participants were *not* required to complete trials correctly before proceeding.

### Results and Discussion

The percentages of correct responses from the second exposure in each of the three phases were grouped according to the six relation types and are presented in Figure 2.

The results of Experiment 2 confirmed the previous findings and indicated again that participants initially produced higher levels of accuracy on specified relations compared with unspecified relations. The unspecified-mixed and unspecified-mixed transitive relations were again the weakest of all performances. However, the feedback intervention employed here appeared to generate considerable improvements on the initially weak performances,

particularly between Phases 1 and 2. Specifically, the mean accuracy of the three unspecified relations together ranged from 53% in Phase 1 to 91% in Phase 2, and 96% in Phase 3.
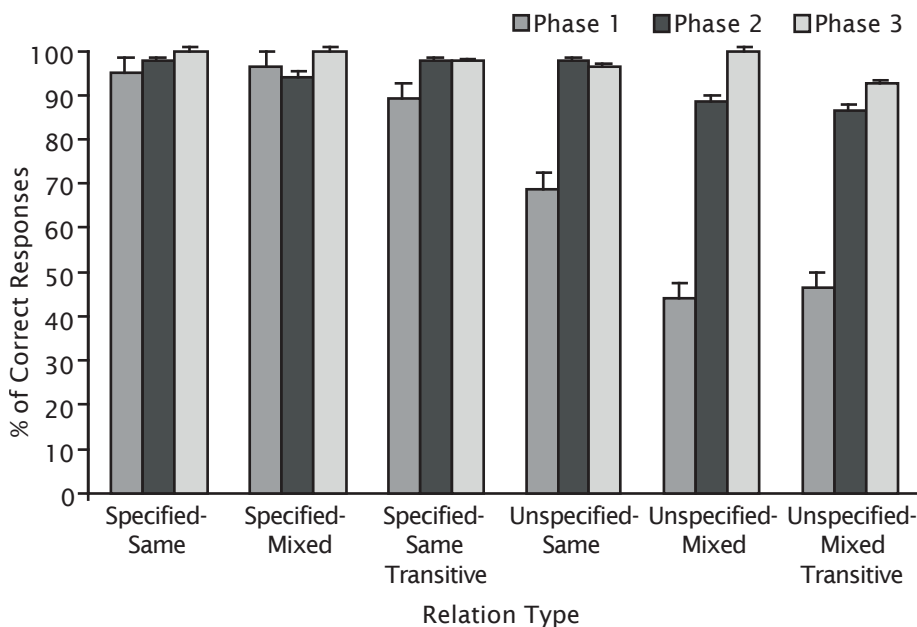


*Figure 2.* The percentage of correct responses on the six relation types presented in the second test exposure of Phases 1, 2, and 3 of Experiment 2.

A 6 × 3 repeated measures ANOVA was again conducted on the data with relation type and phase as between-participant variables. This analysis yielded significant main effects for both relation type, $F(5, 45) = 18.884$, $p < .0001$, $\eta_p^2 = 0.677$, and phase $F(2, 18) = 49.079$, $p < .0001$, $\eta_p^2 = 0.845$, and a significant interaction effect between the two variables $F(10, 90) = 16.925$, $p < .0001$, $\eta_p^2 = 0.653$. Fisher's post hoc analyses once again indicated that the majority of significant differences between relation types were recorded between specified and unspecified relations (see Table 3), but only limited differences remained by Phase 3. Indeed, the only statistical differences that remained in Phase 3 involved the most difficult unspecified-mixed transitive relations, which continued to differ significantly from all other specified relations. These results suggest that although all of the initially weak unspecified relations improved considerably as a result of the written feedback in Phase 2, the unspecified-mixed transitive relations remained significantly weaker than the other relations. As a result of the remaining weakness in the most difficult relation type, Experiment 3 investigated an alternative intervention.

Table 3
*Statistical Comparisons of Participant's Performances on Each Relation Type Across All Three Phases Presented in Experiment 2*

| Relation Type Comparisons | Phase 1 | Phase 2 | Phase 3 |
|---|---|---|---|
| Specified-same vs. Unspecified-same | .0006 | NS | NS |
| Specified-same vs. Unspecified-mixed | <.0001 | .0153 | NS |
| Specified-same vs. Unspecified-mixed transitive | <.0001 | .0022 | .0035 |
| Specified-mixed vs. Unspecified-same | .0004 | NS | NS |
| Specified-mixed vs. Unspecified-mixed | <.0001 | NS | NS |
| Specified-mixed vs. Unspecified-mixed transitive | <.0001 | .0327 | .0035 |
| Specified-mixed transitive vs. Unspecified-same | .0070 | NS | NS |
| Specified-mixed transitive vs. Unspecified-mixed | <.0001 | .0153 | NS |
| Specified-mixed transitive vs. Unspecified-mixed transitive | <.0001 | .0022 | .0431 |
| Unspecified-same vs. Unspecified-mixed | .0010 | .0153 | NS |
| Unspecified-same vs. Unspecified-mixed transitive | .0026 | .0022 | NS |

*Note.* NS = nonsignificant difference.


# Experiment 3

Experiment 3 was designed to investigate the impact of an alternative training procedure, namely, the highlighting of nonarbitrary relations, on test performances—particularly involving weak unspecified relations. Previous research indicated positive outcomes in young children's abilities to derive arbitrary comparative relations as a result of exposure to these relations in nonarbitrary form (Barnes-Holmes et al., 2004) when these relational skills were found to be absent from the children's repertoires. As a result, in the current research, generating a form of nonarbitrary training that might facilitate the target arbitrary comparative relations seemed potentially useful.

This nonarbitrary intervention is based on the RFT assumption that 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. According to Hayes et al. (2001), only in the most artificial of learning environments could one imagine a child responding in accordance with the arbitrary relations of more-than and less-than before first demonstrating the nonarbitrary class of this relational responding. In Experiment 3, therefore, participants were exposed to the experimental trials in nonarbitrary form in Phase 2 in an attempt to investigate the effect of this type of intervention on the subsequent test performances. The nonarbitrary intervention employed

here involved altering the actual sizes of the three coins, so that they were no longer identical, and therefore the target relations were no longer arbitrary.

## Method

*Participants.* A total of 10 naïve participants who were undergraduate students at NUIM took part in Experiment 3. All participants were between 18 and 24 years old; were recruited through class announcements and university notice boards; and did not receive any remuneration for their participation.

*Apparatus.* The automated program employed in Experiment 1 was modified for use in Experiment 3 by changing the relative sizes of the three coins presented during all trials only in Phase 2.

*Procedure.* To accommodate all six different relation types and to be able to present these as nonarbitrary trials, two modifications to the size of the stimuli were necessary: one for specified relations and the other for unspecified relations.

During all trials involving specified relations in Phase 2, the three coins (A, B, and C) were physically different in size: one was small, one was medium, and one was large; and the relative sizes of the coins depended on the trial type. Consider, for example, a specified-same trial in which the participant was instructed as follows: "The blue coin is worth more than the yellow coin and the red coin is worth more than the blue coin." During this trial in Phase 2, the sizes of the coins matched the instruction, such that the blue coin was medium in size, the yellow coin was small, and the red coin was large.

In trials containing unspecified relations, a correct response always involved placing two of the coins (whose location in the jars could not be determined) into the "I cannot know" tin, and placing the third coin into either the large or the small jar (but not in the medium-sized jar). The actual sizes of the coins presented during nonarbitrary trials involving unspecified relations varied across trials. That is, when the coin whose size was specified was worth less than the other two, the other coins may have been both medium or large in size. Alternatively, if the specified coin was worth more, the other two coins may have both been medium or small in size. To present this trial in nonarbitrary form, the two unspecified coins were presented as identical in size and the third coin was different. Consider, for example, an unspecified-same trial in which the participant was instructed as follows: "The blue coin is worth less than the yellow coin and the blue coin is worth less than the red coin." In this case, the red and the yellow coins may have been medium in size on the screen and the blue coin would have been small.

In all other regards, the procedure and the instructions were identical to those in Experiment 1.

## Results and Discussion

The percentages of correct responses from the second exposure in each of the three phases of Experiment 3 were grouped according to the six types of relations, and these are presented in Figure 3.
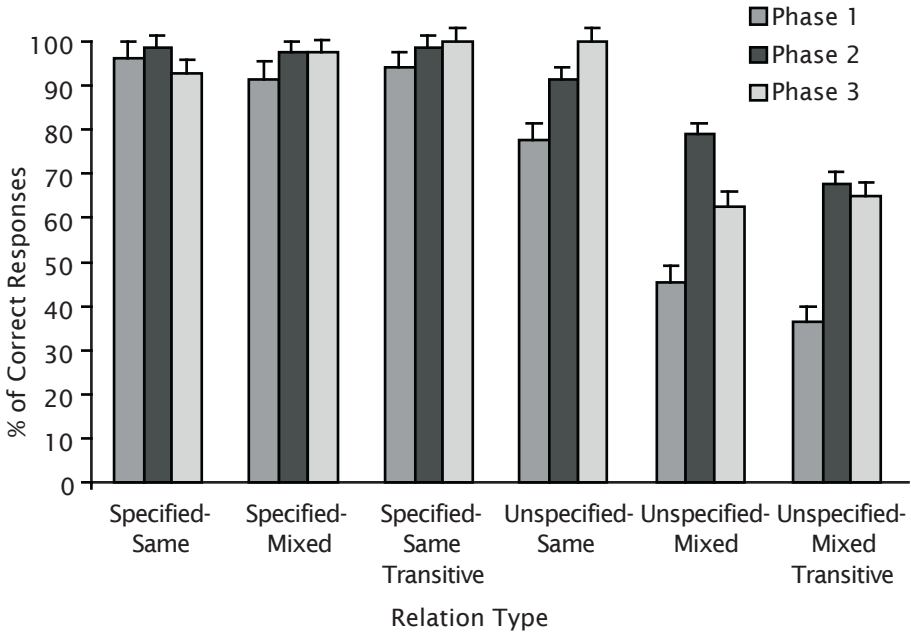
*Figure 3.* The percentage of correct responses on the six relation types presented in the second test exposure of Phases 1, 2, and 3 of Experiment 3.

The results of Experiment 3 confirmed once again initially higher rates of accuracy on specified relations than on unspecified relations overall. Once again, the intervention in Phase 2 appeared to generate improvements on the weak unspecified relations. Specifically, the mean accuracy of unspecified-mixed and unspecified-mixed transitive relations ranged from 45.4% and 36.4% in Phase 1; 78.9% and 67.8% in Phase 2; and 62.7% and 65.2% in Phase 3, respectively. These outcomes suggested some limited improvement in the weakest relations as a result of the nonarbitrary trials in Phase 2.

A $6 \times 3$ repeated measures ANOVA was conducted on the data and again yielded significant main effects for both relation type, $F(5,45) = 26.894$, $p < .0001$, $\eta_p^2 = 0.749$, and phase, $F(2,18) = 13.540$, $p = .0003$, $\eta_p^2 = 0.6$, and a significant interaction effect between the two variables, $F(10,90) = 3.684$, $p = .0004$, $\eta_p^2 = 0.29$. Fisher's post hoc analyses once again indicated that the majority of significant differences between relation types were recorded between specified and unspecified relations (see Table 4). For Experiment 3, the majority of these significant differences remained in Phase 3. These results suggest that although all of the initially weak unspecified relations improved as a result of the nonarbitrary intervention in Phase 2, this improvement was not as notable as that achieved in Experiment 2. Furthermore, the gains achieved in Phase 2 of Experiment 3 were not retained in Phase 3. As a result of the remaining weakness in the most difficult relation type, Experiment 4 investigated an alternative intervention.

Table 4
*Statistical Comparisons of Participant's Performances on Each Relation Type Across All Three Phases Presented in Experiment 3*

| Relation Type Comparisons | Phase 1 | Phase 2 | Phase 3 |
|---|---|---|---|
| Specified-same vs. Unspecified-same | .0118 | NS | NS |
| Specified-same vs. Unspecified-mixed | <.0001 | .0033 | .0003 |
| Specified-same vs. Unspecified-mixed transitive | <.0001 | <.0001 | .0008 |
| Specified-mixed vs. Unspecified-mixed | <.0001 | .0057 | <.0001 |
| Specified-mixed vs. Unspecified-mixed transitive | <.0001 | <.0001 | .0001 |
| Specified-mixed transitive vs. Unspecified-same | .0271 | NS | NS |
| Specified-mixed transitive vs. Unspecified-mixed | <.0001 | .0033 | <.0001 |
| Specified-mixed transitive vs. Unspecified-mixed transitive | <.0001 | <.0001 | <.0001 |
| Unspecified-same vs. Unspecified-mixed | <.0001 | NS | <.0001 |
| Unspecified-same vs. Unspecified-mixed transitive | <.0001 | .0006 | <.0001 |

*Note.* NS = nonsignificant difference.

One possible criticism of the nonarbitrary trials provided in Experiment 3 is that they were simply too few to have a substantive effect on the unspecified relations. To address this issue, Experiment 4 included increased numbers of nonarbitrary trials to determine if this intervention would generate stronger performances on the weak relations.

## Experiment 4

Experiment 4 was designed to increase the number of exposures participants received to nonarbitrary trials. In this experiment, therefore, *all* of the trials across all 6 exposures were presented in nonarbitrary form.

### Method

*Participants.* A total of 10 naïve individuals participated in Experiment 4. Once again, they were undergraduate students aged 18-24, they were recruited through faculty announcements and notice boards, and they did not receive any remuneration for their participation in the experiment.

*Apparatus.* The apparatus used in Experiment 4 was identical to that used in Phase 2 of Experiment 3 except that all of the trials in all three experimental phases were presented in nonarbitrary form.

*Procedure.* The procedure and instructions employed in Experiment 4 were identical to those used previously in that participants received a total of 6 test exposures across three phases and the same instructions were issued. In this experiment, however, all of the trials in the three phases were presented in nonarbitrary form.

### Results and Discussion

Figure 4 presents the percentage of correct responses on the six relation-types from the second exposure in each of the three phases.
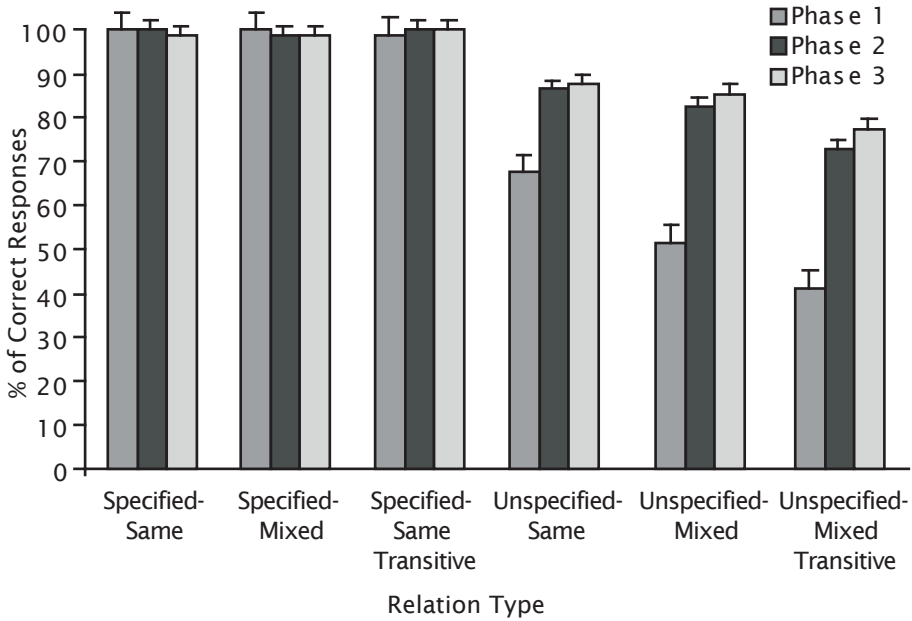
*Figure 4.* The percentage of correct responses on the six relation types presented in the second test exposure of Phases 1, 2, and 3 of Experiment 4.

Figure 4 shows the characteristic pattern of differences between specified and unspecified relations, with higher levels of accuracy overall on the former compared with the latter. The figure shows substantive improvements in the three types of unspecified relations in Phase 2 compared with Phase 1 (i.e., 54% mean accuracy in Phase 1 compared with 80% in Phase 2), and maintenance of these improvements in Phase 3 (83% accuracy).

A $6 \times 3$ repeated measures ANOVA again yielded significant main effects for both relation type, $F(5,45) = 37.339$, $p < .0001$, $\eta_p^2 = 0.806$, and phase, $F(2,18) = 22.774$, $p < .0001$, $\eta_p^2 = 0.717$, and a significant interaction effect between the two variables, $F(10,90) = 3.750$, $p = .0003$, $\eta_p^2 = 0.294$. Fisher's post hoc analyses indicated that the majority of significant differences between relation types were recorded between specified and unspecified relations, and that some of these differences were maintained across phases. Specifically, many significant differences occurred in Phase 1, and these remained in Phase 2 (see Table 5). By Phase 3, some of the differences had gone, but all of those that remained involved unspecified-mixed and unspecified-mixed transitive relations. These results suggest that the initially weak unspecified relations improved in Phase 2, although not considerably, with limited further improvement in Phase 3.

The findings from Experiment 4 indicated that although the presentation of all experimental trials in nonarbitrary form produced considerable improvements on all three types of unspecified relations and these improvements were retained, both types of unspecified-mixed relations

Table 5
*Statistical Comparisons of Participant's Performances on Each Relation Type Across All Three Phases Presented in Experiment 4*

| Relation Type Comparisons | Phase 1 | Phase 2 | Phase 3 |
|---|---|---|---|
| Specified-same vs. Unspecified-same | <.0001 | .0308 | NS |
| Specified-same vs. Unspecified-mixed | <.0001 | .0062 | .0462 |
| Specified-same vs. Unspecified-mixed transitive | <.0001 | <.0001 | .0026 |
| Specified-mixed vs. Unspecified-same | <.0001 | .0479 | NS |
| Specified-mixed vs. Unspecified-mixed | <.0001 | .0104 | .0462 |
| Specified-mixed vs. Unspecified-mixed transitive | <.0001 | <.0001 | .0026 |
| Specified-mixed transitive vs. Unspecified-same | .0001 | .0308 | NS |
| Specified-mixed transitive vs. Unspecified-mixed | <.0001 | .0062 | .0308 |
| Specified-mixed transitive vs. Unspecified-mixed transitive | <.0001 | <.0001 | .0016 |
| Unspecified-same vs. Unspecified-mixed | .0054 | NS | NS |
| Unspecified-same vs. Unspecified-mixed transitive | .0002 | .0296 | NS |

*Note.* NS = nonsignificant difference.

remained significantly weaker than specified relations throughout. In summary, therefore, the improvements resulting from the presentation of *all* nonarbitrary trials in Experiment 4 were similar to those recorded with only one phase of nonarbitrary trials in Experiment 3; and they were not of the same magnitude as those recorded for written feedback in Phase 2 of Experiment 2. Experiment 5 investigated the effect of combining nonarbitrary relations and feedback.

# Experiment 5

In light of the strong improvements recorded for feedback in Experiment 2 and the relatively modest improvements associated with the nonarbitrary interventions, it seemed reasonable to question the impact of combining these two interventions to determine whether their integration would enhance or reduce the effects of either intervention alone. Hence, Experiment 5 attempted to investigate the level of improvement that might be achieved by *combining* written feedback and nonarbitrary training. Because little improvement had been recorded in Experiment 4 relative to Experiment 3 (i.e., in the provision of three phases of nonarbitrary trials compared with only one phase), Experiment 5 consisted of one phase that involved nonarbitrary trials (as in Experiment 3) combined with written feedback (as in Experiment 2).

## Method

*Participants.* A total of 10 naïve participants were involved in Experiment 5. Once again, they were undergraduate students aged 18–24, recruited through faculty notice boards and announcements, and did not receive any remuneration.

*Apparatus.* The apparatus used in Experiment 5 was a combination of the programs from Experiments 2 and 3.

*Procedure.* Phase 2 of Experiment 5 was similar to Experiment 2 in that written feedback was presented, and it was also identical to Experiment 3 in that the trials were presented in nonarbitrary form. Phases 1 and 3 of Experiment 5 were identical to Experiment 1, as were all other aspects of the procedure and the instructions.

## Results and Discussion

Figure 5 presents the percentage of correct responses on the 6 relation types from the second exposure in each of the three phases. The figure shows the characteristic pattern of differences between specified and unspecified relations, with higher levels of accuracy overall on specified relations. Once again, substantive improvements in the three types of unspecified relations resulted in Phase 2 compared with Phase 1 (i.e., 51.4% mean accuracy in Phase 1 compared with 98.4% in Phase 2) and maintenance of these improvements in Phase 3 with high accuracy (99.6%).
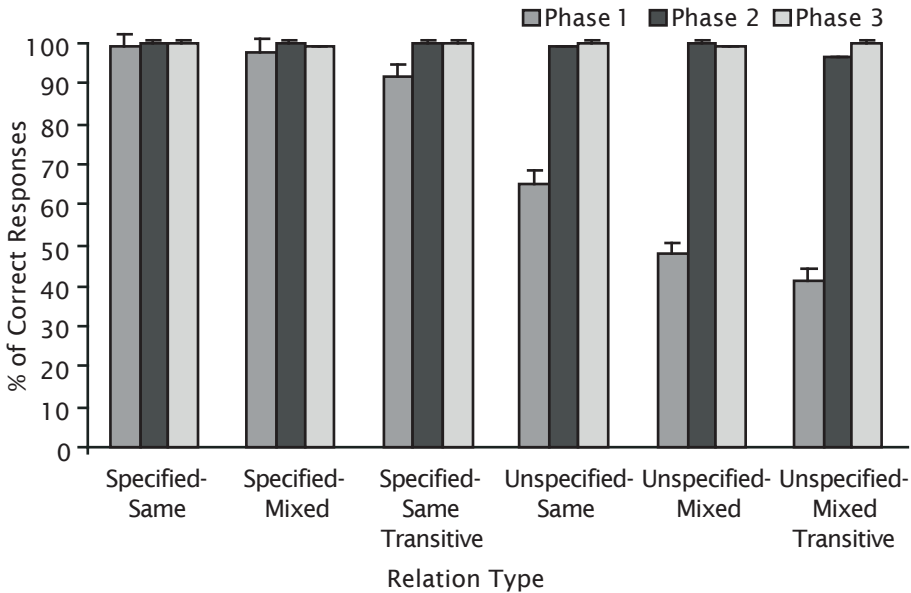


*Figure 5.* The percentage of correct responses on the six relation types presented in the second test exposure of Phases 1, 2, and 3 of Experiment 5.

A $6 \times 3$ repeated measures ANOVA was conducted on the data and yielded significant main effects for both relation types, $F(5,45) = 36.514$, $p < .0001$, $\eta_p^2 = 0.802$, and phase, $F(2,18) = 310.902$, $p < .0001$, $\eta_p^2 = 0.972$, as well as a significant interaction effect between the two variables, $F(10,90) = 34.229$, $p < .0001$, $\eta_p^2 = 0.792$. Fisher's post hoc analyses indicated that the majority of significant differences were recorded between specified and unspecified relations only in Phase 1 (see Table 6). Only three significant differences remained in Phase 2, all of which included unspecified-mixed transitive relations. This pattern of results suggests significant improvements in

unspecified relations in Phase 2 that were maintained in Phase 3. Indeed, performances on unspecified relations in Phases 2 and 3 reached the same levels of accuracy as on the other relations, thus highlighting the efficacy of combining written feedback and nonarbitrary trials. Furthermore, the current data appear to demonstrate stronger improvements than those recorded with either written feedback (Experiment 2) or nonarbitrary trials (Experiment 3) when presented alone.

Table 6
*Statistical Comparisons of Participant's Performances on Each Relation Type Across All Three Phases Presented in Experiment 5*

| Relation Type Comparisons | Phase 1 | Phase 2 | Phase 3 |
|---|---|---|---|
| Specified-same vs. Unspecified-same | <.0001 | NS | NS |
| Specified-same vs. Unspecified-mixed | <.0001 | NS | NS |
| Specified-same vs. Unspecified-mixed transitive | <.0001 | .0077 | NS |
| Specified-mixed vs. Unspecified-same | <.0001 | NS | NS |
| Specified-mixed vs. Unspecified-mixed | <.0001 | NS | NS |
| Specified-mixed vs. Unspecified-mixed transitive | <.0001 | .0077 | NS |
| Specified-mixed transitive vs. Unspecified-same | <.0001 | NS | NS |
| Specified-mixed transitive vs. Unspecified-mixed | <.0001 | NS | NS |
| Specified-mixed transitive vs. Unspecified-mixed transitive | <.0001 | .0077 | NS |
| Unspecified-same vs. Unspecified-mixed | .0054 | NS | NS |
| Unspecified-same vs. Unspecified-mixed transitive | .0002 | NS | NS |

*Note.* NS = nonsignificant difference.

## Experimental Comparisons

To facilitate systematic comparisons of the improvements across the five experiments, the performances on the specified relations were removed and the unspecified performances were grouped for all participants. Figure 6 presents the overall means for each of the five experiments. This figure indicates that the largest improvements occurred in Experiments 2 (feedback) and 5 (feedback and nonarbitrary trials). The improvements in Experiment 4 (all nonarbitrary) were more modest but were greater than those recorded in Experiment 3. Experiment 1 generated the least number of improvements of all five studies.

A $3 \times 3 \times 5$ mixed between ANOVA (with relation type and phase as within-participant variables, and experiment as the between-participant variable) yielded significant main effects for relation type, $F(2,90) = 39.348$, $p < .0001$, $\eta_p^2 = 0.466$, and phase, $F(2,90) = 105.961$, $p < .0001$, $\eta_p^2 = 0.0702$, and experiment, $F(4,45) = 3.855$, $p = .0089$, $\eta_p^2 = 0.255$. Significant interaction effects were also recorded for phase by relation type, $F(4,180) = 4.203$, $p = .0028$, $\eta_p^2 = 0.085$, relation type by experiment, $F(8,90) = 2.929$, $p = .0059$, $\eta_p^2 = 0.207$, and phase by experiment, $F(8,90) = 6.187$, $p < .0001$, $\eta_p^2 = 0.355$. Fisher's post hoc

tests indicated a series of significant differences between the experiments: Experiments 1 vs. 2; 1 vs. 5; 2 vs. 3; 3 vs. 5; and 4 vs. 5. All differences were
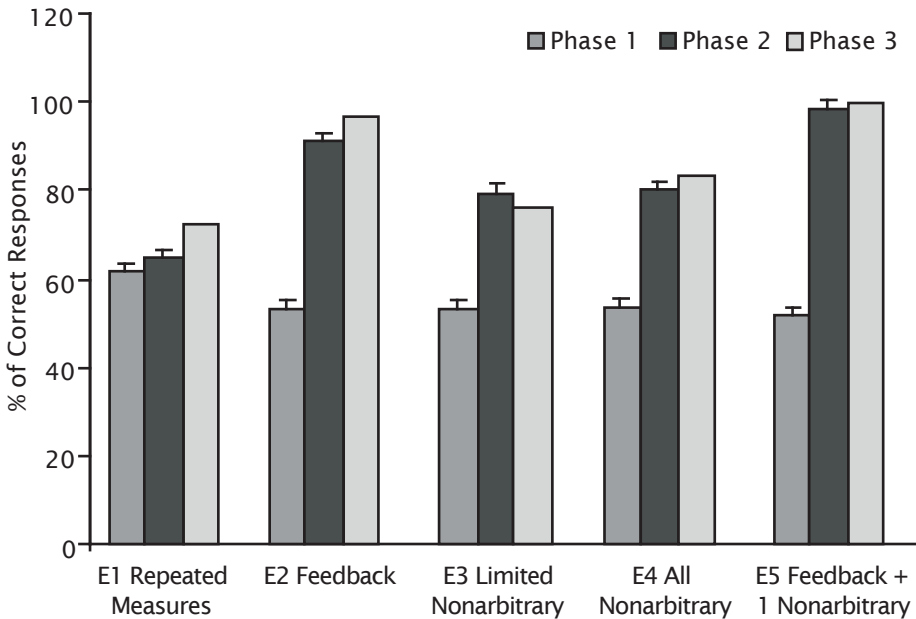


*Figure 6.* A comparison of the percentage of correct responses in each of the five experiments.

significant at the level of $p < .001$. These findings indicate that feedback was significantly better than both repeated exposure and limited nonarbitrary trials. A combination of feedback and nonarbitrary trials was significantly better than repeated exposure, limited nonarbitrary trials, and extended nonarbitrary trials. In all cases, therefore, the provision of feedback resulted in significant improvements, particularly when this feedback was combined with a phase of nonarbitrary trials. However, the combination of feedback and nonarbitrary trials was not significantly better than feedback alone.

## General Discussion

The series of experiments reported here had two primary aims. The first was an attempt to examine a range of more-than and less-than relations that could be readily tested and trained in an experimental context. The second aim was an explicit attempt to develop interventions that would improve target relational performances that were deficient.

The concordance of the data across all five experiments in terms of the superiority of responding on specified versus unspecified relations, as well as the distinct weaknesses in unspecified-mixed and unspecified-mixed transitive relations, was in keeping with the cognitive three-term series literature. Indeed, the experimental outcomes overall were almost identical and provide firm evidence that, in general, adult participants derive specified relations more readily than unspecified; same relations more readily than mixed; and nontransitive relations more readily than transitive.

Although considerable research on the three-term series problem was conducted by cognitive psychologists during the '70s, this period was followed by a dearth in work in this area that was due, at least in part, to conflicting results across studies (Byrne & Handley, 1992; Johnson-Laird & Byrne, 1990; Rips, 1989). Two of the most notable controversies concerned differential outcomes on same versus mixed relations (see Hunter, 1957) and specified versus unspecified relations (Clark, 1969, see also Hayes-Roth & Hayes-Roth, 1975; Moeser & Tarrant, 1977; Sternberg, 1980; see also Byrne & Johnson-Laird, 1989). In attempting to account for the differences across studies, Clark (1969) highlighted the particular difficulties that emerge when, for example, mixed relations and unspecified relations are combined. In support of this suggestion, he reported more errors on unspecified-same than on specified-same relations, and more errors on unspecified-mixed than on specified-mixed relations.

The current data indicated that all participants in all studies produced more errors on unspecified relations than on specified relations overall, particularly when mixed relations were presented. Hence, although specified relations overall appear to be easier to solve than unspecified relations, the level of difference between these relations depends, in part, on whether the unspecified relations contain same or mixed relational terms. These findings are entirely consistent with those reported by Clark (1969).

The distinction between specified and unspecified relations is also consistent with RFT. From this perspective, adult participants come to the experiment with a long history of deriving specified relations, but relatively limited histories of deriving unspecified relations. When unspecified relational difficulties are encountered in the natural language environment, adults are likely to simply seek clarification through additional information. Consider the following example. If, in a conversation, you are told that Mary has two nieces, Josie and Anne, who do not get along very well, then you might ask whether Josie and Anne are sisters, cousins, or no relation in order to understand their relationship more fully and thus the implications of the tensions between them. Indeed, in naturally-occurring unspecified relations, the speaker is likely to specify the relations in order to minimize the ambiguity for the listener (e.g., by saying "I forgot to tell you that they were cousins"). In the current experimental work, the option to receive additional information (i.e., another premise) was not available, and therefore the unspecified relations presented here were particularly unusual, and thus difficult to solve.

In contrast, responding to the specified relations has a much longer and more extensive history, and it forms a large part of standard educational practices from an early age (e.g., "If John has 10 oranges and Peter has 5 oranges, who has more?"). According to RFT, complex repertoires of responding to specified relations are established through multiple exemplars of deriving different types of specified relations across many different contexts, including social and educational contexts (see Barnes-Holmes et al., 2001a). For RFT, then, such differential histories for responding to specified and unspecified relations give rise to very different relational capabilities, as the current evidence clearly reflects.

The current findings also shed light on the conflicting cognitive evidence about same or mixed relational terms. The data here indicated that different performances on same and mixed relations occurred only when

they were presented within specified or unspecified relations. Specifically, performances on specified-same and specified-mixed relations were almost identical, but unspecified-same relations were stronger than unspecified-mixed. Therefore, although the specificity of the relation types appears to play a greater role in determining the ease or difficulty of solving a problem, additional difficulties emerge when mixed relational terms are presented.

This distinction is also consistent with RFT. From this perspective, same relations (i.e., frames of coordination) likely emerge before difference relations (i.e., frames of distinction), because to determine that two things are different, one is likely to have first determined that they are not the same. Indirect support for this view has been obtained from empirical investigations of opposite relations, in which Barnes-Holmes et al. (2001) successfully employed same relations to establish opposite relations when the latter were absent from the children's repertoires. In a similar manner, therefore, two same relations may render a difficult problem easier to solve than mixed relations, particularly in the context of unspecified relations.

The current research contained some conflicting evidence about initial levels of competence on unspecified-same relations and the potential overlap with specified-same relations. Specifically, in all five experiments, responding on specified-same relations was never lower than 85% correct and generally improved to almost perfect responding by Phase 3, irrespective of the intervention. Although responding to unspecified-same relations was considerably better in all cases than performances on the other unspecified relations, some variability occurred on the former across experiments. Indeed, in Experiment 1, responding to these relations was never lower than 82% even in Phase 1; but in Experiment 5, accuracy was only at 65% at the same point in time. In the former case, therefore, accuracy on unspecified-same relations appeared similar to specified-same relations, whereas in the latter cases it did not. However, in all cases where responding on unspecified-same relations was initially low, accuracy improved to almost perfect across phases, irrespective of intervention. Although the variability in responding to unspecified-same relations is difficult to explain, the fact that improvements emerged with repeated exposure alone, and that they were not specific to any one type of intervention, suggests that improvement in these responses would have occurred anyway. As a result, responding here does appear to be more similar than dissimilar to performances on specified-same relations. One might then suggest that the presence of same relations dominated the unspecified features of the task, therefore making these easier to solve than unspecified tasks in which the relations were mixed and similar to same relations that were specified.

Cognitive researchers have also devoted considerable attention to comparisons between performances on transitive versus nontransitive relations. Hayes-Roth and Hayes-Roth (1975), for example, reported that participants took as much as three times longer to solve transitive relations than nontransitive relations (see Russell, McCormack, Robinson & Lillis, 1996, for a similar study with children). However, the results from the current research suggest that transitivity per se may not be a critical variable in determining the difficulty of the problem. Specifically, participants here performed with equally high accuracy on both specified-mixed and specified-mixed transitive, and they performed better on specified-mixed transitive than on unspecified-mixed transitive relations. Taken together, these data suggest

that transitivity is only one feature, and perhaps not the most influential feature, in determining the ease or difficulty with which three-term series problems are resolved. Again, combinations involving unspecified and mixed relations appear to render problems particularly difficult to solve.

A primary aim of the current research was also to examine the utility of a range of interventions that might remediate performances on weak relations. The results from Experiment 1 demonstrated that little or no gains on unspecified relations resulted from a total of 6 repeated exposures to the automated test protocol. Exposure alone, therefore, even across many trials, was not powerful enough to improve the weakest performances. In contrast, a single phase of feedback presented in Experiment 2 facilitated considerable and sustained improvements on all of the weak unspecified relations. Although large improvements were also observed with a single phase of nonarbitrary trials in Experiment 3, weaknesses remained on the unspecified relations. The results from Experiment 4, however, indicated that this shortfall could not be improved by simply increasing the number of exposures to nonarbitrary trials. Indeed, when all of the test trials in Experiment 4 were converted to nonarbitrary trials, the weaknesses in the unspecified relations continued, and as such this intervention offered no advantage over fewer nonarbitrary trials. Experiment 5 examined the potential utility of combining the most successful interventions thus far and demonstrated that feedback was the most active ingredient for improving the weak performances. That is, adding feedback to one phase of nonarbitrary trials produced a better outcome than any amount of nonarbitrary trials alone (i.e., Experiment 5 was better than Experiments 3 or 4), but adding nonarbitrary trials to feedback did not particularly enhance the performances (Experiment 5 was only marginally better than Experiment 2). Hence, although the combination of feedback and nonarbitrary trials was in fact the most successful intervention, and it did result in almost perfect responding on all relations, feedback was clearly a more active ingredient than the inclusion of nonarbitrary trials.

Although the development of interventions for facilitating performances is not a common aim in cognitive research, this approach is almost a hallmark of behavioral psychology. However, although an extensive behavioral literature exists on the role of feedback, only limited evidence exists to date on its role in facilitating derived relational responding (e.g., Healy et al., 2000). The current data support the existing evidence on the positive benefits of feedback on relational responding. Similarly, only limited evidence exists on the role of nonarbitrary training in facilitating arbitrary relational responses (Barnes-Holmes et al., 2001), but again the present findings support existing evidence. No studies to date appear to have examined the combination of these interventions in facilitating derived relational responding, and thus the current evidence is an important first study here.

Taken together, the intervention outcomes compared in the current series of experiments highlighted the utility of feedback in particular in facilitating the various types of comparative relations targeted here. Indeed, the outcomes associated with feedback alone were significantly better than several other interventions, and the combination of feedback with nonarbitrary trials was significantly better than nonarbitrary trials alone (even when the latter accompanied all trials). However, the combination of feedback and nonarbitrary trials was not significantly better than feedback alone. Although feedback is a standard behavioral tool, and as such is not an RFT-specific intervention, its

efficacy is not inconsistent with the theory. As noted previously, RFT would predict that feedback would be effective if it functioned as a type of exemplar training that highlights to the learner all aspects of the target response. This seems likely in the current case in the context of a great many experimental trials, no less than 6 types of relational performance, and an experimental sequence in which trial types were presented randomly. Hence, the feedback appeared to target and reinforce the flexibility necessary to respond correctly across this constantly changing behavioral environment.

The nonarbitrary intervention alone generated notable improvements, particularly when it accompanied all trials (Experiment 4); but this was not *significantly* better than any other intervention, and it was significantly weaker than the combination of nonarbitrary trials and feedback. The nonarbitrary intervention employed here was driven by previous experimental work on establishing arbitrary comparative relations with young children. However, although comparative relations were also targeted in the current research, the differences in the complexity of the relations targeted here may be considerable (including unspecified-mixed transitive relations) and those targeted previously with the children (i.e., these were all specified-same nontransitive relations). That is, in the research with children, the conversion of the arbitrary trials to nonarbitrary trials appeared to simplify the former considerably because the children produced perfect responding on the nonarbitrary trials. In other words, they were already very proficient at the nonarbitrary relations, and these differed only marginally in terms of complexity from the arbitrary relations. Hence, highlighting the relationship between the two meant that training in one type of relation (i.e., nonarbitrary) greatly enhanced training in the other (i.e., arbitrary). In this context, it is not surprising that an intervention in which an arbitrary trial was converted to a nonarbitrary trial would be very effective. Indeed, in other studies where nonarbitrary relations were not at high levels of proficiency, one would not expect such an intervention to be effective; and, indeed, empirical evidence shows that this is the case (Gorham, 2004).

In the current research, however, the nonarbitrary trials were more complex and still required considerable derivation on behalf of participants. That is, although the trials are referred to as nonarbitrary and the items on screen were physically different in size, the nature of the trials in many cases still included unspecified, mixed, and transitive relations. Indeed, even in Experiment 4, where all trials were nonarbitrary, the adult participants did not produce perfect responding. In other words, had these trials been completely nonarbitrary and as such required no derivation, then the assumption would be that adult participants would have made few or no errors. As a result, it was perhaps erroneous to refer to these trials as "nonarbitrary," because in the current experimental context they still had arbitrary features. The outcomes of the nonarbitrary intervention indicated that the highlighting of the nonarbitrary features of the trials in terms of the differences in physical size was facilitative, but did not go far enough because arbitrary features of the trials remained. Indeed, future research might examine whether the separation of the relational features of the tasks or whether the training of the relations in a sequence, for example, from specified to unspecified with increasing complexity, would enhance the impact of this type of intervention.

The current series of experiments was the first to systematically compare

the utility of a range of interventions for facilitating responding to complex comparative relations with adult participants. The range of relations targeted included many not normally present, even in the course of adult discourse, and the interventions ranged from standard behavioral techniques to those more directly driven by RFT. The overlap across the data was considerable and clearly highlighted the utility of the various interventions, particularly on the most complex and weakest relations. The random nature of the experimental task also required high levels of relational flexibility on the part of the adult participants. Taken together, these features of the task and the high level of success associated with some of the interventions point to the potential use of the research here within remedial programs for the establishment of these relations for populations in which they are found to be deficient. In this case, the current work has potential applicability in spite of its abstract nature and complexity.

## References

BARNES-HOLMES, D., HAYES, S.C., DYMOND, S., & O'HORA, D. (2001). Multiple stimulus relations and the transformation of stimulus functions. In S. C. Hayes, D. Barnes-Holmes, & B. T. Roche (Eds.) *Relational Frame Theory: A Post-Skinnerian account of human language and cognition* (pp. 51–72). New York: Plenum.

BARNES-HOLMES, Y., BARNES-HOLMES, D., ROCHE, B., & SMEETS, P.M. (2001a). Exemplar training and a derived transformation of functions 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 functions in accordance with symmetry: II. *The Psychological Record, 51*, 589–604.

BARNES-HOLMES, Y., BARNES-HOLMES, D., SMEETS, P.M., STRAND, P., & FRIMAN, P. (2004). Establishing relational responding in accordance with more-than and less-than as generalized operant behavior in young children. *International Journal of Psychology and Psychological Therapy*, *4,* 531–558.

BYRNE, R.M.J., & JOHNSON-LAIRD, P.N. (1989). Spatial reasoning. *Journal of Memory and Language, 28*, 564–575.

BYRNE, R.M.J., & HANDLEY, S.J. (1992). Reasoning strategies. *Irish Journal of Psychology, 13,* 111–124.

CLARK, H. H. (1969). Linguistic processes in deductive reasoning. *Psychological Review, 76*, 387–404.

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.

GORHAM, M. (2004). Establishing relational responding in accordance with more-than and less-than and transitive relations in normally-developing and young children diagnosed with autism. Unpublished undergraduate thesis of the National University of Ireland, Maynooth.

HAYES-ROTH, B., & HAYES-ROTH, F. (1975). Plasticity in Memorial Networks.

*Journal of Verbal Learning and Verbal Behavior, 14,* 506–522.

HAYES, S.C., BARNES-HOLMES, D., & ROCHE, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition.* New York: Plenum.

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 Experimental Analysis of Behavior, 65,* 182–241.

HORNE, P. J., LOWE, C.F., & HARRIS, F.D.A. (2007). Naming and categorizing in young children: V. Manual sign training. *Journal of Experimental Analysis of Behavior, 87,* 367–381.

HUNTER, I.M.L. (1957). The solving of three term series problems. *British Journal of Psychology, 48,* 286–298.

JOHNSON-LAIRD, P.N., & BYRNE, R.M.J. (1990). Methodological reasoning: Knights, knaves and Rips. *Cognition, 36,* 69–84.

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.

MANI, K., & JOHNSON-LAIRD, P.N. (1982). The mental representation of spatial descriptions. *Memory & Cognition, 10,* 181–187.

MOESER, S.D., & TARRANT, B.L. (1977). Learning a network of comparisons. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 643–659.

RIPS, L.J. (1989). The psychology of knights and knaves. *Cognition, 31,* 85–116.

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.

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.

SIDMAN, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14,* 5–13.

SIDMAN, M. (1994). *Equivalence relations and behavior: A research story.* Boston: Authors Cooperative.

STEELE, D.L., & HAYES, S.C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior, 56,* 519–555.

STERNBERG, R.J. (1980). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General, 109,* 119–159.

STERNBERG, R.J. (1981). Intelligence and nonentrenchment. *Journal of Educational Psychology, 73,* 1–16.

WULFERT, E., & HAYES, S.C. (1988). The transfer of conditional sequencing through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50,* 125–144.