The Psychological Record, 2013, 63, 751–768

CONSTRUCTING RELATIONAL SENTENCES: ESTABLISHING ARBITRARILY APPLICABLE COMPARATIVE RELATIONS WITH THE RELATIONAL COMPLETION PROCEDURE

Anita Munnelly, Gary Freegard, and Simon Dymond

Swansea University

The present study examined the effectiveness of the Relational Completion Procedure (RCP) in establishing arbitrarily applicable comparative relations after training with either linear (Experiment 1) or non-linear (Experiment 2) arbitrary relations. Linear relations were those presented in a sequential order (i.e., A–B, B–C, C–D, etc.), while nonlinear relations were presented in a nonsequential order (i.e., C–D, B–C, A–B, etc.). After nonarbitrary relational training and testing to establish the contextual functions of MORE THAN and LESS THAN, adults underwent constructed-response nonarbitrary relational training and testing with physically dissimilar stimuli in which they were required to "complete" relational sentences in the presence of the cues. Next, during arbitrary relational training with abstract stimuli, participants were randomly assigned to 1 of 3 training groups: All-More (E > D, D > C, C > B, and B > A), All-Less (A < B, B < C, C < D, and D < E), and Less-More (A < B, B < C, D > C, and E > D). Finally, during arbitrary relational testing, probes for mutually entailed relations were followed by probes for 1- and 2-node combinatorially entailed relations. The yield, or percentage of participants passing, was 83% in Experiment 1 and 91.6% in Experiment 2. Our findings demonstrate that the RCP was successful in establishing arbitrarily applicable comparative relations with adults regardless of the order in which the arbitrary relational training trials were presented.

Key words: arbitrarily applicable comparative relations, More Than, Less Than, Relational Completion Procedure, constructed response, linear, adults

An extensive empirical literature now attests to the fact that human participants may come to respond in accordance with multiple stimulus relations, such as Same and Opposite, Before and After, and More Than and Less Than (e.g., Berens & Hayes, 2007; Dymond & Whelan, 2010; Hyland, O'Hora, Smyth, & Leslie, 2012; Munnelly, Dymond, & Hinton, 2010; for a review, see Dymond, May, Munnelly, & Hoon, 2010). Derived

---

This research was conducted as part of the first author's doctoral research program under the supervision of the third author.

We thank Sara Tepaeru Minster for providing the stimuli used during arbitrary relational training and testing.

Correspondence concerning this article should be addressed to Simon Dymond, Department of Psychology, Swansea University, Singleton Park, Swansea SA2 8PP, United Kingdom; E-mail: s.o.dymond@swansea.ac.uk

DOI:10.11133.j.tpr.2013.63.4.004

comparative relations, such as More Than/Less Than, are involved "whenever one event is responded to in terms of a quantitative or qualitative relation along a specified dimension with another event" (Hayes, Fox, Gifford, Wilson, Barnes-Holmes, & Healy, 2001, p. 36). Comparative relations, such as More Than and Less Than, are first learned with nonarbitrary stimuli differing along a specified physical dimension, such as size or quantity, but may be brought to bear on any arbitrary (physically dissimilar) stimuli, given appropriate contextual cues. That is, relational responding may emerge that is not based on the formal, physical properties between the stimuli but rather the arbitrary relational context in which the stimuli are encountered. For instance, a young child may learn that "X is taller than Y," and when asked, "Which is shorter?" he or she may subsequently respond, "Y," without any further training. According to relational frame theory (Hayes, Fox, et al., 2001; Torneke, 2010), such instances of arbitrarily applicable derived comparative relations are controlled solely by the contextual cues "taller" and "shorter," not by physical cues, and these relations may be applied to any stimuli regardless of their physical properties.

Studies examining derived comparative relations of More Than and Less Than have relied heavily on variants of the matching-to-sample (MTS) protocol. For instance, Reilly, Whelan, and Barnes-Holmes (2005) employed a contextually controlled MTS training and testing paradigm to examine comparative responding among adult participants to a five-member relational network. Reilly et al. compared response latencies to arbitrary test relations in three training groups: All-More, All-Less, and Less-More. The first phase of the experiment, termed *nonarbitrary relational training*, aimed to establish contextual control for two cues, More Than and Less Than. For example, participants were presented with one and two basketballs as comparison stimuli in the presence of the MORE THAN contextual cue, where correct selections of two basketballs were reinforced. Nonarbitrary relational testing, which involved the presentation of novel stimulus sets in the absence of feedback, followed. Participants were then exposed to arbitrary relational training involving abstract visual images (i.e., with stimuli that were unrelated to one another along a consistent physical dimension) with the MORE THAN and LESS THAN contextual cues. Participants in the All-More group were trained on B > A, C > B, D > C, and E > D in the presence of the MORE THAN contextual cue, the All-Less group on A < B, B < C, C < D, and D < E in the presence of the LESS THAN cue, and the Less-More group on A < B and B < C in the presence of the LESS THAN cue and D > C and E > D in the presence of the MORE THAN cue (it is important to note that > and < are used here to denote the contextual cues of "more than" and "less than," respectively. Participants were not exposed to these inequality symbols. Instead, two abstract visual images served as contextual cues). During the arbitrary relational test phase, participants were presented with novel stimulus combinations and the two contextual cues, in the absence of feedback. For instance, mutual entailment was tested, for the All-More group, with presentations of A < B, B < C, C < D, and D < E; for the All-Less group, with B > A, C > B, D > C, and E > D; and for the Less-More group, with B > A, C > B, C < D, and D < E. *Mutual entailment* refers to the derived bidirectionality of stimulus relations, whereas *combinatorial entailment* refers to instances in which two or more relations are combined to make a third relation (Hayes, Barnes-Holmes, & Roche, 2001). For all three groups, one- and two-node combinatorial entailment was tested with presentations of A < C, B < D, C < E, C > A, D > B, and E > C (one-node) and A < D, B < E, D > A, and E > B (two-node). Results demonstrated that response latencies were significantly faster on all test relations for participants in the All-More group, in comparison to the other two groups (see also Munnelly et al., 2010).

The use of selection-based MTS procedures, such as those employed by Reilly et al. (2005) demonstrated the ease with which derived relational responding can be generated in a laboratory setting. Such procedures typically involve individuals pointing to, touching, or looking at a stimulus involved in a conditional relation (Shafer, 1993). With respect to the Reilly et al. study, participants were presented with a contextual cue in the top third of the computer screen and two comparison stimuli in the bottom left and right of the screen. To select the comparison stimulus on the left, participants were required to press the "z" key, and to select the comparison stimulus on the right, participants pressed the "m" key. Participants were exposed to a number of conditional discriminations, where they were trained to select, for example, stimulus B and not stimulus A, in the presence of the MORE THAN cue (i.e., B > A). In another variant of MTS, constructed-response matching-to-sample (CRMTS), participants are required to "construct" the correct comparison stimulus from its individual components (e.g., deSouza, Goyos, Silvares, & Saunders, 2007; Dube, McDonald, McIlvane, & Mackay, 1991). For instance, a participant may be presented with the written word *cat* as a sample stimulus, where selecting the letters *c*, *a*, and *t* from a choice pool would indicate the correct comparison stimulus. Dube et al. (1991) used a CRMTS procedure to establish spelling repertoires in two individuals with developmental delays and found it was effective in establishing generalized identity matching for both individuals.

Although there are many advantages associated with selection-based response systems, some limitations have been noted. For instance, the top-down method of presenting the contextual cue above both comparison stimuli does not reflect the order in which individuals encounter relational stimuli in their everyday environment. Typically, when individuals engage in tasks, such as reading non-Arabic languages (e.g., English), the stimuli are presented (read) in sequence from left to right. Indeed, Mackay and Fields (2009) proposed that nonarbitrary properties, such as the position of events in sequences, are critical for performances on learning tasks. Thus, it may be beneficial for researchers examining the emergence of multiple stimulus relations to incorporate training and testing procedures that are more reflective of real-life experiences. The Relational Completion Procedure (RCP; Dymond & Whelan, 2010) originally developed to examine Same and Opposite relations, may provide one such approach.

With the RCP, stimuli are presented sequentially from left to right. The sequential presentation of stimuli from left to right in the RCP is intended to imitate real-world presentations of words and numbers used in educational tasks, such as reading and math tasks. For example, in the RCP, participants are presented with a sample stimulus, contextual cue, and a blank comparison on the top half of the computer screen, followed by three comparison stimuli on the bottom half of the screen. Participants are required to drag and drop one of the comparison stimuli from the bottom of the screen to the blank comparison on the top of the screen. Participants are then required to evaluate their responses by confirming their selection via one of two confirmatory response buttons at the bottom of the screen, "Finish Trial" or "Start Again." This protocol was employed to train and test for the emergence of nonarbitrary and arbitrary relational responding. Results demonstrated a facilitative effect for the confirmatory response requirement, and a greater number of participants successfully completing the experimental task following training and testing with the RCP, in comparison to those who received an MTS training and testing protocol (Dymond & Whelan, 2010). On this basis, the authors suggested that the RCP may have utility as a novel procedure for training and testing other multiple stimulus relations. Finally, the RCP represents an attempt to move beyond standard MTS procedures and provides participants with the opportunity to engage in a more evaluative form of responding (Hayes & Barnes, 1997), the facilitative effects of which remain to be determined.

The current study sought to examine the potential utility of a variation of the RCP combined with a constructed-response training and testing protocol in establishing arbitrarily applicable comparative relations. Participants were first exposed to nonarbitrary relational training and testing, followed by constructed-response nonarbitrary and arbitrary relational training and testing. During the arbitrary relational training phase, participants were exposed to one of three training designs (All-More, All-Less, and Less-More) used previously in MTS-based research (Munnelly et al., 2010; Reilly et al., 2005).

Regardless of group, all participants received arbitrary relational training trials presented in a linear order. That is, participants in the All-More group were trained on B > A, C > B, D > C, and E > D in the presence of the MORE THAN contextual cue, the All-Less group on A < B, B < C, C < D, and D < E in the presence of the LESS THAN contextual cue, and the Less-More group on A < B and B < C in the presence of the LESS THAN contextual cue and D > C and E > D in the presence of the MORE THAN contextual cue. During arbitrary relational testing, all participants were exposed to relevant probes for mutual entailment and common probes for combinatorial entailment.

Experiment 1

Method

Participants. Twelve students, three male and nine female, ranging in age from 18 to 30 years (M_{age} = 21.42, *SD* = 4.40), were recruited through personal contacts and the psychology subject pool at Swansea University. Participants were randomly assigned to the All-More, All-Less, or Less-More training groups at the outset of the experiment.

Apparatus and settings. The experiment was conducted in an experimental room (2 × 3 m) containing a desk, a chair, and a computer with monitor. All training and test trials were presented on a 16-in display screen by a program written in Visual Basic.NET, which controlled all stimulus presentations and recorded all responses.

Materials and stimuli. Two arbitrary images were selected from the Windings font and employed as contextual cues for MORE THAN and LESS THAN, respectively, during the nonarbitrary and arbitrary relational training and testing phases. Twenty-eight stimulus sets were employed during nonarbitrary relational training and testing, and these consisted of images of varying quantities of particular objects. For the arbitrary relational training and testing phases, five images from the Kanji script (see Figure 1) were used to generate a five-term linear relational network (A–B–C–D–E).

Figure 1. The five Kanji images employed during arbitrary relational training and testing, labeled A, B, C, D, and E (participants were never exposed to these labels).

General Procedure

The general procedural sequence was as follows: Phase 1A: Nonarbitrary Relational Training and Phase 1B: Nonarbitrary Relational Testing; Phase 2A: Constructed-Response Nonarbitrary Relational Training and Phase 2B: Constructed-Response Nonarbitrary Relational Testing; Phase 3: Constructed-Response Arbitrary Relational Training; Phase 4: Arbitrary Relational Test 1; and Phase 5: Arbitrary Relational Test 2.

During all training and testing phases, the computer screen was separated in two; the top two thirds of the screen was blue, while the bottom third was white. During Phase 1, the sample stimulus appeared first in the top left-hand side of the screen. Following a delay of 1 s, the contextual cue (e.g., MORE THAN or LESS THAN) appeared in the upper center of the screen, and a blank yellow square was presented following a 1-s delay in the top right-hand side of the screen. Following a further delay of 1 s, two comparison stimuli appeared simultaneously in the lower third of the screen. The screen position (i.e., left or right) of these comparisons was counterbalanced across trials.

In order to make a response, participants were required to “drag” one of the two comparison stimuli and “drop” it in the blank yellow square. Immediately upon making this selection, a red border appeared around the comparison stimulus, to highlight participants' selections. Participants then clicked the left mouse button and held it down while dragging their selection to the blank yellow square. Releasing the left mouse button allowed the selected comparison to "drop" into the blank yellow square. At the same time, the screen position in which the comparison stimulus originally appeared was replaced by a blank yellow square.

Once the selected comparison was placed in the blank yellow square, two confirmatory response buttons appeared simultaneously at the bottom of the screen. One button was labeled "Finish Trial," and when participants hovered the mouse cursor over this button, a small text box with the caption "Click Here to Finish Trial" appeared onscreen. The second button was labeled "Start Again," and hovering over this button produced the caption "Click Here to Start Again." If the participant pressed the "Start Again" button, this cancelled the selection and resulted in all stimuli returning to their original positions before the selection was made; that is, the comparison stimulus that was selected returned to either the lower left or right portion of the screen, and the blank yellow square returned to the upper right of the screen (see Figure 2).

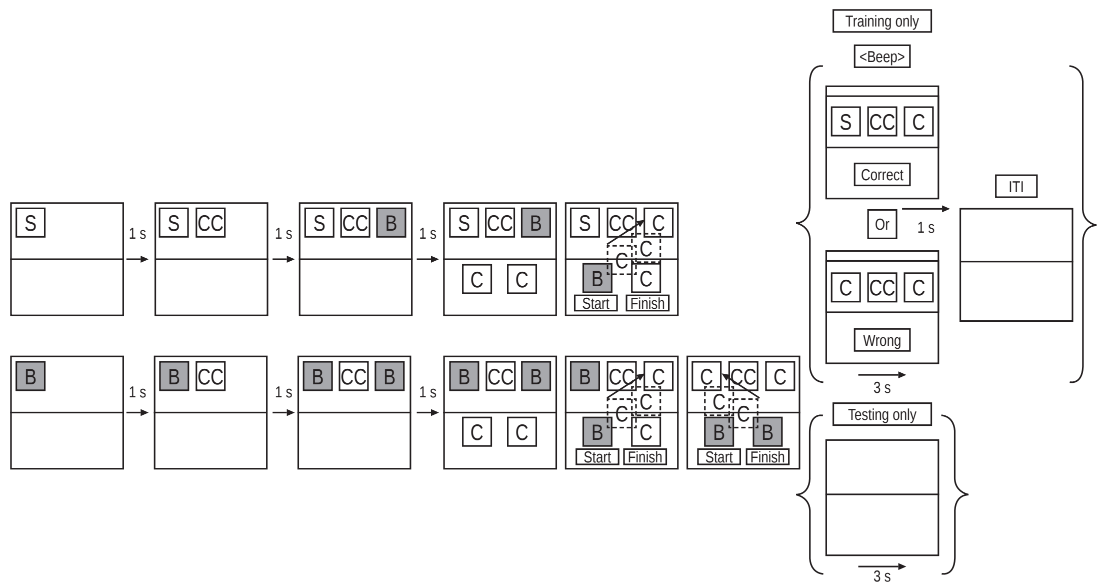

Figure 2. Schematic diagram of the sequence of presentation of stimuli during the nonarbitrary relational training and testing phases (top) and the constructed-response nonarbitrary and arbitrary relational training and test phases (bottom). *Note*. S = sample; CC = contextual cue; B = blank square; C = comparison; ITI = intertrial interval. A dashed line represents dragging of a comparison stimulus. "Finish" and "Start" represent the confirmatory response buttons. Arrows pointing from B to C illustrate that once selected, the comparison stimulus moved to the top portion of the screen, while its original screen position was replaced by a blank square.

All stimuli remained onscreen until the participant pressed the "Finish Trial" button. During training, trials were followed by feedback presented on a blue background for a duration of 3 s. When a participant made a correct response, feedback consisted of, from left to right, the sample, the contextual cue, and the comparison stimulus the participant had selected on the previous trial. A yellow border surrounded all three images, and the word "Correct!" was presented in black underneath. A brief audible beep was presented following the word "Correct!" The only difference between feedback for a correct selection and feedback for an incorrect selection was that the word "Correct!" was now replaced by the word "Wrong" and no audible tone followed this feedback. During all testing trials, no feedback was presented; instead, the screen cleared and remained blue for 3 s. A 2-s intertrial interval (ITI) followed each trial, where the screen cleared and remained blue for the duration of the ITI.

The presentation of stimuli differed during Phases 2 through 5. For example, participants were presented with a blank yellow square followed by a contextual cue and another blank yellow square in the top portion of the screen. Similar to Phase 1, two comparison stimuli were again presented on the lower portion of the screen, but the sample stimulus in the upper left-hand side of the screen was now replaced with a blank yellow square. During these phases, participants were required to "construct" their responses from left to right in the top portion of the computer screen. Participants were instructed to place one of the comparison stimuli in the top left blank yellow square and the other comparison in the top right blank yellow square (see Figure 2). Both the initial response and confirmatory response requirements were identical to Phase 1. Again, all training trials were followed by feedback, whereas feedback was omitted during all test phases.

A task feedback thermometer was displayed in the center right of the screen during all training and testing phases (Fienup, Covey, & Critchfield, 2010). During training, the thermometer displayed the mastery criterion needed to complete training (e.g., "You need this many correct to move on: 10"), the current number of correct responses (e.g., 6 out of 10), which was incremented following every correct response. During testing, the thermometer displayed the total number of trials in the particular test phase and the current trial number, and it was incremented following every response.

Procedure

Phase 1A: Nonarbitrary Relational Training. This training phase aimed to establish contextual control over responding for two arbitrary visual images (MORE THAN and LESS THAN) to comparison stimuli of varying quantities (see Appendix A for participant instructions). For example, on a given trial, participants were presented with a sample (e.g., two basketballs), a contextual cue (e.g., MORE THAN), and a blank yellow square in the top portion of the screen. Two comparison stimuli (e.g., one and four basketballs) were also presented on the bottom portion of the screen. In this instance, placing the comparison stimulus containing one basketball in the blank yellow square counted as a correct response. On the other hand, if two basketballs were again presented as the sample, alongside the contextual cue for LESS THAN, and one and four basketballs as comparison stimuli, placing the comparison stimulus containing four basketballs in the blank yellow square was reinforced. All training trials were followed by feedback presented for 3 s and by an ITI of 2 s.

Four stimulus sets were employed during nonarbitrary relational training. Mastery criterion for this training phase was set at 10 consecutive correct responses. If participants met this criterion, they immediately proceeded to the nonarbitrary relational test phase. However, if participants failed to meet this criterion following exposure to 240 training trials, they were then exposed to a second nonarbitrary relational training phase, which involved four novel stimulus sets.

Phase 1B: Nonarbitrary Relational Testing. This phase was similar to Phase 1A, except that four novel stimulus sets were presented and all feedback was omitted. Participants were presented with a total of eight test trials and were required to respond correctly across all eight trials in order to progress to the next phase of the experiment. However, if participants failed to meet this criterion, they were re-exposed to nonarbitrary relational training (i.e., Phase 1A) involving the same four stimulus sets. This was again followed by a nonarbitrary relational test.

Phase 2A: Constructed-Response Nonarbitrary Relational Training. The purpose of this phase was to train participants to "construct" the relation between two comparison stimuli, in the presence of a particular contextual cue (see Appendix B for participant instructions). On each trial, participants were presented with a blank yellow square, a contextual cue, and another blank yellow square in the top portion of the screen. Participants were also presented with two comparison stimuli on the bottom portion of the screen. For example, participants were presented with the contextual cue for MORE

THAN and six and four guitars as the comparison stimuli. A correct response consisted of dragging and dropping the six guitars in the top left blank yellow square and the four guitars in the top right blank yellow square, *in that sequence*. Similarly, if six guitars and four guitars were presented as comparison stimuli in the presence of the LESS THAN contextual cue, placing the four guitars in the top left blank yellow square and the six guitars in the top right blank yellow square was reinforced. Again, feedback was presented following all training trials.

Participants were presented with four stimulus sets during training, and mastery criterion was set at 10 consecutive correct responses. If participants successfully met training criterion, they were immediately exposed to the constructed-response nonarbitrary relational test phase. However, similar to Phase 1A, if participants were unsuccessful in meeting the training criterion after exposure to 240 training trials, they were re-exposed to constructed-response nonarbitrary relational training, followed again by constructed-response nonarbitrary relational testing.

Phase 2B: Constructed-Response Nonarbitrary Relational Testing. This phase was identical to Phase 2A, with the exception that participants were presented with four novel stimulus sets, and feedback was no longer provided. Participants were exposed to eight test trials and were required to respond correctly across all trials to progress to the next phase. If participants failed to meet this criterion, they were re-exposed to constructed-response nonarbitrary relational training involving the same four stimulus sets. This was again followed by a constructed-response nonarbitrary relational test phase.

Phase 3: Constructed-Response Arbitrary Relational Training. This phase commenced immediately upon completion of Phase 2B (see Appendix C for participant instructions). Similar to Phase 2, participants were presented with a blank yellow square, a contextual cue, and another blank yellow square in the top portion of the screen. Again, two comparison stimuli were presented simultaneously on the bottom portion of the screen. However, during this phase, the comparison stimuli consisted of arbitrary images (i.e., Kanji script), which are labeled, for the purposes of clarity, A, B, C, D, and E (see Figure 1).

Participants were presented with training trials in a linear order, and training pairs differed between the All-More, All-Less, and Less-More training groups. The All-More group were trained with B > A, C > B, D > C, and E > D trials in the presence of the MORE THAN contextual cue; the All-Less group with A < B, B < C, C < D, and D < E trials in the presence of the LESS THAN contextual cue; and the Less-More group with A < B and B < C trials in the presence of the LESS THAN contextual cue, and D > C and E > D trials in the presence of the MORE THAN contextual cue. All training pairs were presented in this order for all three groups (see Table 1).

For all groups, the four training pairs were presented a total of three times each, resulting in a block of 12 training trials. Mastery criterion for the arbitrary relational training phase was set at 12 out of 12 correct responses on any given block. Training blocks were repeated until participants achieved criterion.

Phase 4: Arbitrary Relational Test 1. Upon reaching the arbitrary relational training criterion, participants were exposed to an arbitrary relational test phase that probed for the properties of mutual entailment alongside maintenance of the baseline arbitrary training relations. All feedback was now omitted, and participants were presented with 8 test trials, each presented 4 times, which resulted in a total of 32 test trials (see Table 1). Dependent on the training group, the mutually entailed test trials presented to participants differed between the groups. For example, the mutually entailed relations presented to the All-More group were A < B, B < C, C < D, and D < E; to the All-Less group, B > A, C > B, D > C, and E > D; and to the Less-More group, B > A, C > B, C < D, and D < E.

Mastery criterion for this phase was set at a minimum mean of 12 out of 16 correct responses on the baseline relations. For the mutually entailed relations, participants were required to make three out of four correct responses on each individual mutually entailed test trial. If participants were successful in meeting criterion on both the baseline and mutually entailed relations, they progressed to a second arbitrary relational test phase. However, if participants failed to reach this mastery criterion, they were re-exposed to the experimental task from the very beginning for a maximum of three further exposures.

Table 1
Training and Test Trials Received by the Three Groups in Experiments 1 and 2

Group	Relation type	Test trial type					
		Phase 4: Arbitrary Relational Test 1					
		Specific relations in each group					
All-Less	Baseline	A < B	B < C	C < D	D < E		
	ME	B > A	C > B	D > C	E > D		
All-More	Baseline	B > A	C > B	D > C	E > D		
	ME	A < B	B < C	C < D	D < E		
Less-More	Baseline	A < B	B < C	D > C	E > D		
	ME	B > A	C > B	C < D	D < E		
		Phase 5: Arbitrary Relational Test 2					
		Specific relations in each group					
All-Less	Baseline	A < B	B < C	C < D	D < E		
All-More	Baseline	B > A	C > B	D > C	E > D		
Less-More	Baseline	A < B	B < C	D > C	E > D		
		Relations common to all groups					
	CE1	C > A	D > B	E > C	A < C	B < D	C < E
	CE2	D > A	E > B	A < D	B < E		

Note. *Baseline* refers to test trials involving directly trained relations presented in the absence of feedback. ME = mutually entailed; CE1 and CE2 = one- and two-node combinatorially entailed relations, respectively. The inequality symbols, < (less than) and > (more than), denote the contextual cue presented and indicate which comparison should be selected over the other, with the reinforced comparison shown on the left and the unreinforced comparison on the right. It is important to note that the actual contextual cues used in the present study consisted of abstract visual images and not the inequality symbols described here, which are used for the purposes of clarity.

Phase 5: Arbitrary Relational Test 2. This test phase commenced immediately on completion of Phase 4. Participants were presented with probes for one- and two-node combinatorially entailed relations, along with the four baseline relations. Each trial was presented 4 times in a test block of 56 trials in a quasi-random order. Participants were required to make a minimum of 12 out of 16 correct responses on the baseline relation trials. The same one- and two-node combinatorially entailed relations were presented to all groups (Table 1). Fields, Verhave, and Fath (1984) defined a *node* as a stimulus linked by training with at least two other stimuli. For example, when probing for emergent relations such as BD, the C stimulus is the node separating the two stimuli. The number of nodes that separate the stimuli between which an emergent relation is formed can vary (e.g., one-, two-, and three-node relations) and is referred to as *nodal distance*. Thus, in the current study, participants were trained and tested on a five-term relational network, and were exposed to probes for 6 one-node relations (e.g., D > B, with C as the node) and 4 two-node relations (e.g., E > B, with C and D as nodes; see Table 1). Across these probe trials, participants were required to make a minimum of three out of four correct responses on each individual one- and two-node test trial. If participants were successful in meeting the Phase 5 criteria, the experiment ended. However, if this criterion was not met, participants were re-exposed to the entire sequence from Phase 1 for a maximum of three further exposures.

Results and Discussion

Of the 12 participants who took part in Experiment 1, 2 (P2 and P3) failed to achieve the predetermined Phase 4 arbitrary relational test (mutual entailment) criterion within the maximum four exposures to testing (see Table 2) and were excused from further participation. Ten participants successfully completed the experiment, with the total number of test exposures required ranging from 1 to 4 (M = 1.5, SD = 0.85; see Table 2).

Table 2
Individual Data and Trials to Criterion for Participants in Experiment 1 During Nonarbitrary and Arbitrary Relational Training

Participant	Phases 1A & 2A: NARB & CR-NARB Relational Training		Phase 3: CR-Arbitrary Relational Training	Phase 4: Arbitrary Relational Test 1 (mutual entailment)		Phase 5: Arbitrary Relational Test 2 (combinatorial entailment)		
				Baseline	ME	Baseline	CE1	CE2
All-Less								
1	10	10	24	16/16	16/16	16/16	24/24	16/16
2[a]	12	10	132	8/16	11/16			
	10	10	12	8/16	12/16			
	13	10	12	7/16	8/16			
	10	10	12	10/16	10/16			
3[a]	22	10	132	8/16	9/16			
	15	10	24	10/16	4/16			
	10	10	60	7/16	10/16			
	10	10	12	6/16	7/16			
4	41	10	48	16/16	16/16	16/16	23/24	16/16
All-More								
5	12	10	48	3/16	0/16			
	10	10	24	16/16	15/16	14/16	16/24	14/16
	20	10	12	16/16	16/16	16/16	24/24	16/16
6	22	10	36	16/16	16/16	16/16	24/24	16/16
7	13	10	36	16/16	16/16	15/16	19/24	16/16
	13	10	12	16/16	16/16	16/16	24/24	16/16
8	15	10	60	16/16	16/16	16/16	24/24	16/16
Less-More								
9	40	10	132	16/16	16/16	16/16	24/24	16/16
10	17	10	60	16/16	15/16	16/16	24/24	16/16
11	12	10	48	16/16	16/16	16/16	23/24	16/16
12	10	17	36	16/16	15/16	16/16	14/24	16/16
	14	12	12	16/16	15/16	16/16	20/24	16/16
	15	10	12	16/16	16/16	16/16	24/24	16/16

Note. *Baseline* refers to test trials involving directly trained relations. NARB = nonarbitrary; CR = constructed response; ME = mutually entailed; CE1 and CE2 = one- and two-node combinatorially entailed relations, respectively. Data are displayed for the number of correct responses to the baseline and mutually entailed relations during Test 1 and to the baseline and one- and two-node relations during Test 2.
[a]Participant failed to complete the experiment.

Table 2 displays the trials to criterion and participants' number of correct responses to the baseline and mutually entailed relations during Test 1, and to the baseline and one- and two-node relations during Test 2. Results demonstrate that the 10 participants who passed the experimental task displayed high levels of accuracy on the baseline, mutually entailed, and combinatorially entailed relations during the critical arbitrary relational test phases. There was little difference between the three groups' accuracy during testing. It may, therefore, be concluded that the RCP training and testing protocol was successful in establishing arbitrarily applicable comparative relations of More Than and Less Than.

Experiment 2

The arbitrary relational training tasks in Experiment 1 were presented in a linear, sequential order. Previous research has suggested that the order in which the training pairs are presented may have an effect on test performance (Hunter, 1957; Russell, McCormack, Robinson, & Lillis, 1996). More specifically, Hunter (1957) argued that performances at test are weakened when training pairs are presented in a non-linear order (i.e., randomly) compared to a linear order (i.e., sequentially). For instance, Russell et al. (1996) proposed that transitive test pairs, such as B < D and A < D, are easier to solve (i.e., select B over D and A over D) when training pairs are presented in a linear order (e.g., A < B, B < C, C < D, and D < E) compared to a nonlinear order (e.g., B < C, D < E, A < B, and C < D).

Two previous studies conducted under the rubric of RFT have investigated this issue (Gorham, Barnes-Holmes, Barnes-Holmes, & Berens, 2009; Vitale, Barnes-Holmes, Barnes-Holmes, & Campbell, 2008). Gorham et al. (2009) exposed five typically developing children and three children with autism to More Than and Less Than arbitrary relations involving two, three, four, or five identically sized coins. All children initially failed baseline tests involving the targeted (A–B, B–C, A–B–C, and A–B–C–D) and transitive relations (B–D) and thus were exposed to a training phase, in which these relations were presented in a linear order. Results from both experiments demonstrated that seven participants met training criterion and passed a subsequent test phase for the targeted and transitive relations when they were presented with novel stimulus sets. In addition, Vitale et al. (2008) examined the effects of linearity on arbitrary comparative test performances with a three-term series task. According to the definition of *linearity* adopted in the current study, Vitale et al. exposed participants to tasks involving specified-same (linear; A > B and B > C or A < B and B < C) and specified-same transitive (nonlinear; A > B and C > A or A < B and C < A) arbitrary relational training relations. In addition, participants were exposed to a number of other training tasks involving unspecified relations (e.g., A > B and C > B). Irrespective of whether the trials were presented in a linear or nonlinear order, all participants were exposed to the same arbitrary relational test. Across five experiments, the authors found that accuracy was comparably high on three-term linear and nonlinear arbitrary relational tests.

In Experiment 2, therefore, we sought to replicate and extend the findings of Experiment 1 by presenting the arbitrary relational training trials in a nonlinear order.

Method

Participants. Twelve students, seven male and five female, ranging in age from 19 to 34 years (M_{age} = 22.58, *SD* = 4.34), were recruited via an email announcement and the psychology subject pool at Swansea University. Participants were randomly assigned to the All-More, All-Less, or Less-More training groups at the outset of the experiment.

Procedure. The procedure for Experiment 2 was identical to that of Experiment 1, with the exception of the presentation of tasks in the arbitrary relational training phase (Phase 3). In this phase, the training pairs were presented in a nonlinear order. That is, in each group, the computer program quasi-randomly presented each of the four arbitrary relational training trials (Table 1), with the constraint that each trial could not appear more than twice consecutively. Mastery criterion for this training phase and all other phases remained the same.

Results and Discussion

Of the 12 participants who took part in Experiment 2, 1 participant (P5) failed to meet criterion on the arbitrary relational test, within four exposures to testing. However, 11 participants successfully completed both arbitrary relational test phases, requiring from one to three (M = 1.45, SD = 0.69) exposures to testing to do so (see Table 3).

Table 3
Individual Data and Trials to Criterion for Participants in Experiment 2 During Nonarbitrary and Arbitrary Relational Training

Participant	Phases 1A & 2A: NARB & CR-NARB Relational Training		Phase 3: CR-Arbitrary Relational Training	Phase 4: Arbitrary Relational Test 1 (mutual entailment)		Phase 5: Arbitrary Relational Test 2 (combinatorial entailment)		
				Baseline	ME	Baseline	CE1	CE2
All-Less								
1	17	10	48	15/16	15/16	12/16	24/24	15/16
2	11	10	36	16/16	16/16	16/16	24/24	16/16
3	12	10	24	16/16	16/16	16/16	24/24	15/16
4	19	10	60	13/16	16/16	16/16	23/24	16/16
All-More								
5[a]	10	11	108	16/16	16/16	16/16	9/24	8/16
	15	10	12	16/16	14/16	15/16	8/24	8/16
	10	10	12	16/16	16/16	16/16	8/24	9/16
	10	10	12	15/16	16/16	16/16	9/24	8/16
6	10	10	48	15/16	16/16	16/16	24/24	16/16
7	13	10	72	16/16	15/16	16/16	24/24	16/16
8	12	10	72	13/16	16/16	16/16	8/24	10/16
	10	10	12	16/16	16/16	15/16	24/24	16/16
Less-More								
9	26	10	72	15/16	16/16	15/16	24/24	16/16
10	12	10	36	15/16	16/16	16/16	8/24	0/16
	10	10	12	16/16	16/16	16/16	7/24	1/16
	10	10	12	16/16	16/16	16/16	24/24	16/16
11	15	10	84	16/16	16/16	16/16	19/24	15/16
	10	10	12	16/16	16/16	16/16	21/24	15/16
12	12	10	72	16/16	15/16	16/16	12/24	16/16
	10	10	12	16/16	16/16	16/16	24/24	16/16

Note. *Baseline* refers to test trials involving directly trained relations. NARB = nonarbitrary; CR = constructed response; ME = mutually entailed; CE1 and CE2 = one- and two-node combinatorially entailed relations, respectively. Data are displayed for the number of correct responses to the baseline and mutually entailed relations during Test 1 and to the baseline and one- and two-node relations during Test 2.
[a]Participant failed to complete the experiment.

Table 3 displays the trials to criterion and number of correct responses participants made to the baseline and mutually entailed relations during Test 1, and to the baseline and one- and two-node relations during Test 2. Results demonstrated that the 11 participants who passed the experimental task displayed high levels of accuracy on the baseline, mutually entailed, and combinatorially entailed relations during the critical arbitrary relational test phases. There was little difference between the three groups' accuracy during testing. Thus, it may be concluded that the RCP was successful in establishing responding in accordance with the relational frames of More Than and Less Than when the arbitrary relations were trained in a non-linear sequence.

Further Analyses

Additional analyses were conducted across both experiments to examine the overall yield (i.e., the number of participants who successfully passed both arbitrary relational tests) and the effectiveness of two arbitrary training designs (i.e., linear and nonlinear) in facilitating the emergence of derived comparative relations.

Experiments 1 and 2

In Experiments 1 and 2, there were yields of 83% (10 out of 12 participants) and 91.6% (11 out of 12 participants), respectively. Thus, the RCP with training pairs presented in a nonlinear format was found to have a slight advantage over the RCP with training pairs presented in a linear order, in terms of the overall yield. This difference was not statistically significant ($p = .55$).

Arbitrary Training

The mean number of trials required by participants in Experiment 1 to complete arbitrary relational training was 83, whereas participants in Experiment 2 required a mean of 97 trials. This difference was not statistically significant ($p = .80$).

General Discussion

The present findings demonstrate that the RCP is an effective protocol for establishing derived comparative relations in adults and, in so doing, extend the findings of earlier studies on the relational frames of Same and Opposite (Dymond, Ng, & Whelan, 2013; Dymond, Roche, Forsyth, Whelan, & Rhoden, 2007, 2008; Dymond & Whelan, 2010). In what follows, we will address some of the potential factors responsible for the present performances.

The linearity, or sequential presentation, of arbitrary relational training pairs varied across the two experiments. Previous research conducted in domains outside behavior analysis suggests that the order in which arbitrary relational training pairs are presented may influence arbitrary test performance (e.g., Hunter, 1957; Russell et al., 1996). Cognitive models of such performance differences suggest that presenting the training pairs in a linear order allows participants to covertly organize and arrange the stimuli into a unified linear representation or mental model, which is then inspected during testing to make inferences about the novel, nonadjacent stimulus pairs (e.g., Acuna, Sanes, & Donoghue, 2002; Sedek & von Hecker, 2004; Williams, Avery, Woolard, & Heckers, 2012). Behavior analysts, however, eschew such mediational explanations and have instead sought to develop parsimonious accounts of the effects of linearity on performance during tests for derived comparative relations (Gorham et al., 2009; Munnelly et al., 2010; Vitale et al., 2008).

To our knowledge, the present study is the first to formally compare the performance of adults on linear (Experiment 1) and non-linear (Experiment 2) arbitrary relational training tasks using a five-member network and three training schedules (All-More, All-Less, and Less-More). Reilly et al. (2005) only reported response latencies to derived comparative test relations (i.e., performance accuracy was not assessed), and Whelan, Barnes-Holmes, and Dymond (2006) first established a seven-member relational network (A–B–C–D–E–F–G) in order to examine subsequent transformation of functions. Therefore, the present study extends Vitale et al.'s (2008) research examining the effects of linearity on responses to the relational frame of comparison with adult participants, from a three- to a five-member relational network. Moreover, the current findings suggest that the linearity of training pairs may not be a critical factor in the ability to respond in accordance with the relational frame of comparison.

The comparably high, accurate test performances observed in the present experiments may be partially explained by the response requirements of the RCP. In comparison with the majority of MTS procedures, the RCP allows participants to first complete the relation and then either evaluate it (by confirming their selection) or initiate a new selection (Dymond & Whelan, 2010). In a traditional MTS task, participants select a comparison by clicking on it in the presence of a particular contextual cue (Dymond & Barnes, 1995; Munnelly et al., 2010; Reilly et al., 2005). Dymond and Whelan (2010) compared the effectiveness of the RCP and MTS in establishing Same and Opposite relations and found that a greater number of participants successfully passed the experiment following exposure to the RCP protocol. In addition, the authors found a facilitative effect for the confirmatory response requirement, with a greater number of participants completing the experiment when they were provided with the opportunity to evaluate and confirm their responses in comparison to those who were not. However, the current study did not undertake such an investigation and, therefore, future research should seek to manipulate the presence and absence of the confirmatory response requirement.

The influence exerted by the sequential presentation of stimuli from left to right in the current study highlights the importance of stimulus sequences in experimental tasks. As the RCP was developed to mimic the verbal relational processes involved in everyday tasks, such as reading and sentence completion, a comparison of performances on a present version of the task with a right-to-left version (as seen in other languages, such as Arabic) may help to identify the potential facilitative role of this sequence (Dymond & Whelan, 2010). This issue warrants further empirical attention.

Our findings may have important implications for future studies seeking to examine the emergence of relational reasoning abilities and transitive inference (TI) in both human and nonhuman populations. For example, an ordinal relation may emerge between stimuli A and C (e.g., A+C−, where + and − represent the reinforced and non-reinforced responses, respectively) once a relation has been established between stimuli A and B (A+B−) and between B and C (B+C−; Vasconcelos, 2008). Indeed, numerous studies have found evidence for the expression of TI in young children and nonhumans when minimal or no instructions were provided (e.g., Bryant & Trabasso, 1971; Lazareva, Smirnova, Bagozkaja, Zorina, Rayevsky, & Wasserman, 2004; McGonigle & Chalmers, 1977). One such study, conducted by McGonigle and Chalmers (1977), involved training squirrel monkeys on the relation between four adjacent stimulus pairs (A+B−, B+C−, C+D−, and D+E−), followed by tests involving non-adjacent stimulus pairs (e.g., BD). In this study, no verbal instructions were employed and the stimuli consisted of weighted cans consisting of different colors. In effect, McGonigle and Chalmers first trained the squirrel monkeys to associate weight with color, which led to successful selections of B over D in the BD pair during the critical inferential test phase (e.g., B was lighter than D). With respect to the current model, if all verbal instructions were omitted, it may be possible to examine the emergence of derived comparative responding in nonhumans. Nonhumans are capable of learning to respond to nonarbitrary relations (e.g., Harmon, Strong, & Pasnak, 1982; Reese, 1968), but evidence for derived relations in nonhumans is currently equivocal at best (Dymond, Roche, & Barnes-Holmes, 2003; Hayes, 1989). Indeed, if the current model was successful in establishing arbitrary comparative responding in nonhumans, it would have important implications for theoretical accounts of derived relations (Hayes, Fox, et al., 2001) and for the development of procedures for use with individuals that lack sophisticated verbal repertoires (Berens & Hayes, 2007).

During nonarbitrary and arbitrary phases of the current study, a constructed-response protocol was adopted based on findings suggesting that this method of training and testing is effective in generating identity relations and equivalence relations (e.g., deSouza et al., 2007; Dube et al., 1991; Stromer & Mackay, 1992a, 1992b). It is difficult, however, to assess whether or not the act of constructing responses facilitated the effects of the evaluative component of the RCP because participants were exposed to a phase involving nonarbitrary relational training and testing *before* the constructed-response phases (Barnes-Holmes, Hayes, Dymond, & O'Hora, 2001). Therefore, future studies examining the effectiveness of the constructed-response protocol in facilitating the emergence of derived relational responding should first expose participants only to the constructed-response protocol.

Across both experiments, only three participants failed to meet criterion in the arbitrary relational test. Of these, two participants failed to meet criterion on the mutually entailed relations (Experiment 1), whereas the remaining participant failed to meet criterion on the combinatorially entailed relations (Experiment 2). In previous studies on derived comparative relations, when the predicted relations fail to emerge, participants were exposed to exemplar training involving nonarbitrary relations. For instance, Vitale et al. (2008) employed a multiple-exemplar training intervention, which consisted of converting the difficult arbitrary trials into a nonarbitrary form. That is, the arbitrary stimuli, which consisted of three coins, were altered so that they were no longer identical in size and thus no longer arbitrary. Vitale et al. reported that the intervention produced considerable improvements in arbitrary relational test performances (see also Berens & Hayes, 2007). Therefore, it may have been beneficial for the small proportion of participants in the current experiments who failed to meet the arbitrary relational test criterion to have been administered a similar nonarbitrary-based intervention.

A noteworthy feature of the current study was the inclusion of a "task feedback thermometer" during training and testing phases. The thermometer incremented following correct responses, and because the training criterion involved responding consecutively across a block of trials, it reset to 0 when an error was made. During testing phases, the thermometer incremented following each response (correct and incorrect) and was not reset if participants made an error. This additional onscreen feedback was employed during all phases of the experiments as a motivating operation (Michael, 1993) to increase engagement with the task and to make phase progression and task termination reinforcing. Further research should seek to determine the relative effectiveness of the task thermometer feedback during training and testing of multiple stimulus relations.

The alternative training and testing protocol employed in the current experiments may be suited to a touchscreen-based response system. For example, the use of touch-sensitive screen monitors may maximize responding, as participants are presented with fewer response options, which in turn may help minimize errors. Incorporating touchscreen technology with the RCP may also be advantageous for use with applied populations (e.g., Dube et al., 1991). Moreover, the current procedures may be suited to the use of topography-based response systems in establishing different relational frames (e.g., Lipkens & Hayes, 2009).

Finally, it may be possible to adapt the current protocol to examine the facilitative effects of a relational training and testing intervention that relied solely on experimenter-delivered verbal instructions. For instance, if participants were exposed to an extensive history of exemplar training across multiple stimulus sets, in which they were explicitly informed of the relative value of each stimulus, it may be possible to examine the emergence of derived comparative responding with the presentation of untrained, novel stimulus combinations at testing. Indeed, previous studies have found such interventions to be effective in generating symmetrical and comparative responding in both typically developing children and children with autism, when these repertoires were found to be deficient (e.g., Barnes-Holmes, Barnes-Holmes, Roche, & Smeets, 2001a, 2001b; Berens & Hayes, 2007; Gorham et al., 2009). Thus, future research should seek to examine the effectiveness of the RCP in establishing derived comparative responding with such populations.

References

ACUNA, B. D., SANES, J. N., & DONOGHUE, J. P. (2002). Cognitive mechanisms of transitive inference. *Experimental Brain Research, 146*, 1–10.

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*, 589–603.

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, NY: Plenum.

BERENS, N. M., & HAYES, S. C. (2007). Arbitrarily applicable comparative relations: Experimental evidence for a relational operant. *Journal of Applied Behavior Analysis, 40*, 45–71.

BRYANT, P. E., & TRABASSO, T. (1971). Transitive inferences and memory in young children. *Nature, 232*, 456–458.

DESOUZA, S., GOYOS, C., SILVARES, E. F. M., & SAUNDERS, R. R. (2007). Emergence of printing and spelling skills from constructed-response matching-to-sample instruction. *European Journal of Behavior Analysis, 8*, 49–64.

DUBE, W. V., MCDONALD, S. J., MCILVANE, W. J., & MACKAY, H. A. (1991). Constructed-response matching to sample and spelling instruction. *Journal of Applied Behavior Analysis, 24*, 305–317.

DYMOND, S., & BARNES, D. (1995). A transfer of self-discrimination functions in accordance with the derived stimulus relations of sameness, more-than, and less-than. *Journal of the Experimental Analysis of Behavior, 64*, 163–184.

DYMOND, S., MAY, R. J., MUNNELLY, A., & HOON, A. E. (2010). Evaluating the evidence base for relational frame theory: A citation analysis. *The Behavior Analyst, 33*, 97–117.

DYMOND, S., NG, T. C., & WHELAN, R. (2013). Establishing arbitrarily applicable relations of same and opposite with the relational completion procedure: Selection-based feedback. *The Psychological Record, 63*, 111–130.

DYMOND, S., ROCHE, B., & BARNES-HOLMES, D. (2003). The continuity strategy, human behavior, and behavior analysis. *The Psychological Record, 53*, 333–347.

DYMOND, S., ROCHE, B., FORSYTH, J. P., WHELAN, R., & RHODEN, J. (2007). Transformation of avoidance functions in accordance with the relational frames of same and opposite. *Journal of the Experimental Analysis of Behavior, 88*, 249–262.

DYMOND, S., ROCHE, B., FORSYTH, J. P., WHELAN, R., & RHODEN, J. (2008). Derived avoidance learning: Transformation of avoidance response functions in accordance with the relational frames of same and opposite. *The Psychological Record, 58*, 271–288.

DYMOND, S., & WHELAN, R. (2010). Derived relational responding: A comparison of match-to-sample and the relational completion procedure. *Journal of the Experimental Analysis of Behavior, 94*, 37–55.

FIELDS, L., VERHAVE, T., & FATH, S. J. (1984). Stimulus equivalence and transitive associations: A methodological analysis. *Journal of the Experimental Analysis of Behavior, 42*, 143–157.

FIENUP, D. M., COVEY, D. P., & CRITCHFIELD, T. S. (2010). Teaching brain-behavior relations economically with stimulus equivalence technology. *Journal of Applied Behavior Analysis, 43*, 19–33.

GORHAM, M., BARNES-HOLMES, Y., BARNES-HOLMES, D., & BERENS, N. (2009). Derived comparative and transitive relations in young children with and without autism. *The Psychological Record, 59*, 221–246.

HARMON, K., STRONG, R., & PASNAK, R. (1982). Relational responses in tests of transposition with rhesus monkeys. *Learning and Motivation, 13*(4), 495–504.

HAYES, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 51*, 385–392.

HAYES, S. C., & BARNES, D. (1997). Analyzing derived stimulus relations requires more than the concept of stimulus class. *Journal of the Experimental Analysis of Behavior, 68*, 235–244.

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

HAYES, S. C., FOX, E., GIFFORD, E. V., WILSON, K. G., BARNES-HOLMES, D., & HEALY, O. (2001). Derived relational responding as learned behavior. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 21–49). New York, NY: Plenum.

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

HYLAND, J., O'HORA, D., SMYTH, S., & LESLIE, J. C. (2012). Sequential responding in accordance with temporal relational cues: A comparison of before and after. *The Psychological Record, 62*, 463–484.

LAZAREVA, O. F., SMIRNOVA, A. A., BAGOZKAJA, M. S., ZORINA, Z. A., RAYEVSKY, V. V., & WASSERMAN, E. A. (2004). Transitive responding in hooded crows requires linearly ordered stimuli. *Journal of the Experimental Analysis of Behavior, 82*, 1–19.

LIPKENS, R., & HAYES, S. C. (2009). Producing and recognizing analogical relations. *Journal of the Experimental Analysis of Behavior, 91*, 105–126.

MACKAY, H. A., & FIELDS, L. (2009). Syntax, grammatical transformation and productivity: A synthesis of stimulus sequences, equivalence classes and contextual control. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding: Applications for learners with autism and other developmental disabilities* (pp. 209–235). Oakland, CA: New Harbinger Publications.

MCGONIGLE, B. D., & CHALMERS, M. (1977). Are monkeys logical? *Nature, 267*, 694–696.

MICHAEL, J. (1993). Establishing operations. *The Behavior Analyst, 16*, 191–206.

MUNNELLY, A., DYMOND, S., & HINTON, E. C. (2010). Relational reasoning with derived comparative relations: A novel model of transitive inference. *Behavioural Processes, 85*, 8–17.

REESE, H. W. (1968). *The perception of stimulus relations: Discrimination learning and transposition*. New York, NY: Academic Press.

REILLY, T., WHELAN, R., & BARNES-HOLMES, D. (2005). The effect of training structure on the latency of responses to a five-term linear chain. *The Psychological Record, 55*, 233–249.

RUSSELL, J., MCCORMACK, T., ROBINSON, J., & LILLIS, S. G. (1996). Logical (versus associative) performance on transitive reasoning tasks by children: Implications for the status of animals' performance. *Quarterly Journal of Experimental Psychology, 49B*, 231–244.

SEDEK, G., & VON HECKER, U. (2004). Effects of subclinical depression and aging on generative reasoning about linear orders: Same or different processing limitations? *Journal of Experimental Psychology: General, 133*, 237–260.

SHAFER, E. (1993). Teaching topography-based and selection-based verbal behavior to developmentally disabled individuals: Some considerations. *The Analysis of Verbal Behavior, 11*, 117–133.

STROMER, R., & MACKAY, H. A. (1992a). Delayed constructed-response identity matching improves the spelling performance of children with mental retardation. *Journal of Behavioral Education, 2*, 139–156.

STROMER, R., & MACKAY, H. A. (1992b). Spelling and emergent picture-printed word relations established with delayed identity matching to complex samples. *Journal of Applied Behavior Analysis, 25*, 893–304.

TORNEKE, N. (2010). *Learning RFT: An introduction to relational frame theory and its clinical applications*. Oakland, CA: New Harbinger Publications.

VASCONCELOS, M. (2008). Transitive inference in nonhuman animals: An empirical and theoretical analysis. *Behavioural Processes, 78*, 313–334.

VITALE, A., BARNES-HOLMES, Y., BARNES-HOLMES, D., & CAMPBELL, C. (2008). Facilitating responding in accordance with the relational frame of comparison: Systematic empirical analyses. *The Psychological Record, 58*, 365–390.

WHELAN, R., BARNES-HOLMES, D., & DYMOND, S. (2006). The transformation of consequential functions in accordance with the relational frames of more-than and less-than. *Journal of the Experimental Analysis of Behavior, 86*, 317–335.

WILLIAMS, L. E., AVERY, S. N., WOOLARD, A. A., & HECKERS, S. (2012). Intact relational memory and normal hippocampal structure in the early stage of psychosis. *Biological Psychiatry, 71*, 105–113.

Appendix A

Instructions for Phase 1A: Nonarbitrary Relational Training

Thank you for agreeing to participate in this study. You will be presented with a series of images on the top half of the screen from left to right. Then you will be presented with 2 images on the bottom of the screen. Your task is to observe the images that appear from left to right and place one of these images from the bottom in the blank, yellow square. To select the image on the bottom, click on it once, and to place it in the blank square, click on this once. To confirm your choice, click “Finish Trial.” If you wish to make another choice, then click “Start Again.” Sometimes you will receive feedback on your choices, but at other times you will not. Your aim is to get as many tasks correct as possible. It is always possible to get a task correct, even if you are not given feedback. If you have any questions, please ask the experimenter. Please press the OK button below to begin the experiment!

Appendix B

Instructions for Phase 2A: Constructed-Response Nonarbitrary Relational Training

The first phase of the experiment is now finished. You will now be presented with two blank yellow squares in the top left- and right-hand sides of the screen and one image in the center top of the screen. Then you will be presented with two images on the bottom of the screen. Your task is to drag and drop one image at a time from the bottom of the screen into the blank yellow squares. You must drag and drop an image into the left-hand blank yellow square and then drag and drop the next image into the right-hand yellow square. To select the image on the bottom, click on it once, and to place it in the blank square, click on this once. To confirm your choice, click “Finish Trial.” If you wish to make another choice, then click “Start Again.” Sometimes you will receive feedback on your choices, but at other times you will not. Your aim is to get as many tasks correct as possible. It is always possible to get a task correct, even if you are not given feedback. If you have any questions, please ask the experimenter. Please press the OK button below to continue!

Appendix C

Instructions for Phase 3: Constructed-Response Arbitrary Relational Training

The second phase of the experiment is now finished. You will again be presented with two blank yellow squares in the top left- and right-hand sides of the screen and one image in the center top of the screen. Then you will be presented with two images on the bottom of the screen. Your task is to drag and drop one image at a time from the bottom of the screen into the blank yellow squares. You must drag and drop an image into the left-hand blank yellow square and then drag and drop the next image into the right-hand yellow square. To select the image on the bottom, click on it once, and to place it in the blank square, click on this once. To confirm your choice, click “Finish Trial.” If you wish to make another choice, then click “Start Again.” Sometimes you will receive feedback on your choices, but at other times you will not. Your aim is to get as many tasks correct as possible. It is always possible to get a task correct, even if you are not given feedback. If you have any questions, please ask the experimenter. Please press the OK button below to continue!