The Psychological Record, 2009, 59, 221–246

DERIVED COMPARATIVE AND TRANSITIVE RELATIONS IN YOUNG CHILDREN WITH AND WITHOUT AUTISM

Marie Gorham, Yvonne Barnes-Holmes, and Dermot Barnes-Holmes

National University of Ireland, Maynooth

Nicholas Berens

Center for Advanced Learning, Reno

The current study comprised 2 experiments to generate relational responding in typically developing children and children with autism. In Study 1, the children were exposed to a problem-solving task that involved the presentation of 2, 3, or 4 identically sized coins to test and train the arbitrary relations of more than and less than. All 8 children failed baseline tests involving 4 coins and were exposed to training of the A-B, B-C, A-B-C, and A-B-C-D relations. Seven proceeded rapidly through training and passed tests with a novel set. They then participated in Study 2. Four typically developing children and 2 with autism failed baseline B-D tests. Although the former proceeded rapidly through B-D training with 4 and 5 coins, the latter required interventions. All 7 children thereafter demonstrated B-D relations with 4 and 5 coins. These findings support relational frame theory and its use in educational interventions with developmentally delayed populations.

The emergence of novel or untrained behavior has captured the attention of behavioral researchers since the 1970s because of its obvious implications for educational training (e.g., Sidman, 1971). Put simply, the emergence of educationally significant skills and the generalization of these skills to novel contexts may substantively alter, and even reduce, the demands of direct instruction, at least in the long term. The behavioral literature contains numerous attempts to do this, including pivotal responding regimes (Koegel, Carter, & Koegel, 2003), stimulus equivalence procedures (Carr & Felce, 2000), and other programs based on multiple stimulus relations (McHugh, Barnes-Holmes, & Barnes-Holmes, 2004).

Numerous behavioral researchers have argued that complex human verbal behavior is comprised of repertoires of derived relational responding, and this basic tenet underpins the theoretical approaches that comprise both stimulus equivalence (Sidman, 1971) and relational frame theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001). The primary difference between the two

Correspondence concerning this article should be addressed to Yvonne Barnes-Holmes, Department of Psychology, National University of Ireland, Maynooth, Maynooth, Co. Kildare, Ireland. E-mail: Yvonne.Barnes-Holmes@nuim.ie

approaches is that the latter incorporates a broad range of arbitrary relations beyond equivalence or coordination, including comparison, difference, opposition, hierarchy, and perspective taking.

As an operant theory of language and cognition, RFT argues that verbal behavior that cannot be traced back to a history of direct contingency control emerges instead from a history of reinforced multiple exemplars. Imagine a child who learns that a ship is bigger than a boat and then responds that a boat is smaller than a ship, without being explicitly trained to do so. According to RFT, this ability to derive a smaller-than relation from a bigger-than relation occurs because the child has been trained to do so across other relevant exemplars using, for instance, direct reinforcement and imitation (e.g., if A > B, then B < A; if C > D, then D < C; if E > F, then F < E; and so on). In this sense, the derived smaller-than relation (between boat and ship) is not genuinely novel but is instead part of a generalized operant response class.

A similar operant explanation also applies to more complex relational responses, in which more than two elements are involved. For example, a direct history of relevant exemplar training likely would be required for a child to derive a smaller-than relation involving three, rather than two, elements (e.g., if a ship is bigger than a boat and a boat is bigger than an oar, then an oar is smaller than a ship). That is, the child would have learned previously across numerous exemplars that if A > B and B > C, then A > C and C < A.

According to RFT, generalized relational operants may differ in their levels of flexibility. Consider a child who has learned across exemplars to derive four-element relations in which training A > B, B > C, and C > D gives rise to A > D and D < A. Once again, the derived performance observed is not genuinely novel because it has been explicitly trained across previous exemplars. Imagine now, however, that the child is asked about the relationship between B and D and responds appropriately (B > D and D < B). If this B-D relation had not been trained in previous four-element exemplars, this may indicate that the operant of combining more-than/less-than relations has itself generalized (Hayes & Wilson, 1996). In other words, relational performances that have not been specifically trained across exemplars but are nonetheless relationally coherent with that training will be observed. At this point, the relational operant possesses a degree of flexibility that extends beyond the specific relational patterns that were trained across the exemplars. Such flexibility is deemed to be a critically important property of human language and cognition (Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Hayes & Wilson, 1996).

Only three recent studies have reported research that has directly addressed the foregoing analysis of novel verbal behavior (Barnes-Holmes, Barnes-Holmes, Smeets, Strand, & Friman, 2004; Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Berens & Hayes, 2007). In the first study, 3 children, ages 4 to 6 years, were exposed to a basic problem-solving task that involved identically sized paper coins presented in an attempt to test and train responding in accordance with the relations of more than and less than. On each trial, the researcher described how the coins compared to one another in terms of their value and instructed the child to pick the coin that would “buy as many sweets as possible.” For example, on a trial involving three coins, the child was told that coin A buys more than coin B and that coin B buys more than coin C. The child was then asked to select the coin that buys the most (i.e., coin A). Throughout the study, several trial types were designed to test the more-than and less-than relations among three coins (e.g., A > B > C; A < B < C; C > B > A; and C < B < A), and multiple stimulus sets were employed as exemplars to train the target relations. All 3 children failed to pass the initial tests but demonstrated the relational responses after a history of training with exemplars.

In the second study (Barnes-Holmes, Barnes-Holmes, & Smeets, 2004), the same researchers employed the same basic procedure to test and train opposition relations in another group of young children for whom the target relations were again found to be absent.

In the third study, Berens and Hayes (2007) replicated and extended the procedure reported by Barnes-Holmes, Barnes-Holmes, Smeets, Strand, et al. (2004) with 4 typically developing children (approximate age 4 years). Specifically, these researchers employed lengthier baseline testing, nonlinear trial types, and a multiple-probe design with trained and untrained stimuli. Again, although all 4 children failed the baseline tests (some failed many tests), they all demonstrated the target relations after training and, most importantly, generalized across trial types (i.e., trained on linear but generalized to nonlinear).

Although all three studies were conducted with typically developing children, the researchers highlighted the potential utility of their training protocol for developmentally delayed populations whose verbal behavior would remain significantly impaired in the absence of these key repertoires of derived relational responding (e.g., those diagnosed with autism spectrum disorders [ASD]).

Study 1 of the current research attempted to replicate and extend Barnes-Holmes, Barnes-Holmes, Smeets, Strand, et al. (2004) and Berens and Hayes (2007) by establishing repertoires of responding in accordance with more than and less than in samples of typically developing children and children with autism. In addition, the previous research also was extended by the use of four coins instead of three during the testing and training sequences. This latter extension, in Study 2 reported here, facilitated the investigation of relational responses that had not been trained across the previous exemplars. Consider a child who had been trained that A > B, B > C, and C > D and is asked to choose the coin that buys the most from among all four coins. The child could then be presented with a novel set of coins with the same relations (E > F, F > G, and G > H) but asked to pick between only two coins (e.g., F and H). If the child chooses the correct coin without a history of exemplar training for choosing that coin, this would provide evidence for a flexible, generalized relational operant of more than/less than.

Study 1

Method

Participants

A total of 8 children participated in Study 1, 5 typically developing and 3 independently diagnosed with ASD. The typically developing children, 3 males (Participants 1, 2, and 4) and 2 females (Participants 3 and 5), were between 6 years, 5 months and 8 years, 11 months in age. All 5 children attended the same classroom in a mainstream school in Mayo, Ireland. They were selected on the basis of their age and the fact that neither their parents nor their school teachers had identified them as presenting any form of learning difficulty. The children with autism (Participants 6, 7, and 8) were all male and between 6 years and 9 years, 4 months of age. These children attended the CABAS (Comprehensive Application of Behavior Analysis to Schooling) educational facility in Dublin, Ireland. All 3 children were categorized as speakers and early readers according to the CABAS model. A speaker can respond appropriately to spoken verbal behavior and is able to vocally direct others. An early reader can read and write simple words.

Setting and Materials

Each session was conducted in a small, quiet classroom in the two respective schools. The maximum duration of a session was 20 min, and the children were exposed to no more than four sessions per week. Each child participated individually. Throughout the trials, the researcher and the participant were seated side by side at a small table. The stimuli consisted of 24 laminated paper circles referred to throughout the study as "coins." All 24 coins were identical in size; 6 were blue, 6 red, 6 green, and 6 yellow. The coins were divided into six stimulus sets (i.e., Sets 1–6, with 4 coins in each set), with a coin of each color in each set. Although the coins were identical in size and some also were identical in color, each individual coin depicted a specific visual feature (e.g., two crossed lines or a single diagonal line). The four stimuli contained within each set were designated as A, B, C, and D, but participants were not aware of these labels. Stimulus sets were never mixed.

Several additional materials were employed as reinforcers for correct responding and on-task behavior (e.g., strawberries and candies). For the 4 typically developing children, a single reinforcer (a chocolate bar) was provided at the end of each session for overall participation and compliance (i.e., they did not gain access to the edible reinforcer on the basis of correct responding). For the children with ASD, reinforcers were provided for correct responding, and these included candies, strawberries, stickers, toys, and access to play activities. The schedules of reinforcement for compliant and/or correct responding implemented for these children differed for each child in order to ensure consistency with their individualized learning programs.

Relational Repertoires and Trial Types

Study 1 established four basic repertoires of relational responding for each of the 7 children. These repertoires consisted of A-B, B-C, A-B-C, and A-B-C-D relations. There were four trial types (two more than and two less than) presented for each of the four relational repertoires, and these are presented in Table 1 (each line represents a trial type). At the beginning of each trial that contained A-B relations both coins were placed side by side on the table. The researcher then said, "We are going to play a birthday game. I want you to imagine that it is your birthday today, and you have to go to the shops to get sweets for your birthday party. If this coin (researcher pointed to the coin on the left) buys more sweets than this coin (researcher pointed to the coin on the right), which would you take to buy as many sweets as possible?" In order to respond correctly, each child was required to select the one coin that would buy the most sweets (in this example, the coin on the left).

Table 1
The Four Basic Target Relational Repertoires and Constituent Trial Types Presented in Study 1 and an Additional Repertoire (A-B-C-D-E Relations) Conducted With P5 Only

A-B Relations								
A	Less than →	B*						
A*	Less than ←	B						
A*	More than →	B						
A	More than ←	B*						
B-C Relations								
		B	Less than →	C*				
		B*	Less than ←	C				
		B*	More than →	C				
		B	More than ←	C*				
A-B-C Relations								
A	Less than →	B	Less than →	C				
A*	Less than ←	B	Less than ←	C				
A*	More than →	B	More than →	C				
A	More than ←	B	More than ←	C*				
A-B-C-D Relations								
A	Less than →	B	Less than →	C	Less than →	D		
A*	Less than ←	B	Less than ←	C	Less than ←	D		
A*	More than →	B	More than →	C	More than →	D		
A	More than ←	B	More than ←	C	More than ←	D*		
A-B-C-D-E Relations								
A	Less than →	B	Less than →	C	Less than →	D	Less than →	E*
A*	Less than ←	B	Less than ←	C	Less than ←	D	Less than ←	E
A*	More than →	B	More than →	C	More than →	D	More than →	E
A	More than ←	B	More than ←	C	More than ←	D	More than ←	E*

Note. Each line represents a trial type. Arrows indicate direction of researcher pointing. * Indicates a correct response.

Programmed Consequences

A correct response on all training trials consisted of the participant handing the correct coin to the researcher. The researcher then responded with "Yes," "That is correct," "Good girl/boy," or "Have a candy," as appropriate. An incorrect response was defined as making an incorrect choice (i.e., selecting the wrong coin or more than one coin) or emitting no response within 10 s of the start of the trial. If a child emitted an incorrect response, the researcher immediately provided the correct answer by saying, for example, "This coin buys as many sweets as possible" while lifting the correct coin from the table. The next trial was then presented immediately. If a participant made any comments during any trial the researcher simply said, "We can talk about that after we have finished our work." No programmed consequences followed any test trial.

Procedure

At the beginning of the study, each child was first exposed to a baseline test of all four relational repertoires (A-B, B-C, A-B-C, and A-B-C-D; see Table 1). The four trial types contained within each of the four relational repertoires were similar in structure and differed only in terms of whether the targeted relations were more than or less than and whether the researcher pointed from left to right or from right to left at the beginning of the trial. These four trial types contained two more-than and two less-than trials. Of the two more-than trials, one involved the researcher pointing from left to right (e.g., A to B), and the other involved the researcher pointing from right to left (e.g., B to A). The left/right pointing was manipulated in the same way for the less-than trials.

The trial types for the four relational repertoires also differed in terms of the number of coins involved. Specifically, the A-B and B-C trial types involved only two coins, the A-B-C trial types involved three coins, and the A-B-C-D trial types involved four coins. Once again, however, all of these trial types were presented in the same linear sequence outlined in Table 1.

Each baseline test consisted of a total of 32 trials, with two exposures to each of the 16 trial types (see Table 1). These were divided into four separate blocks of 8 quasi-random trials for the A-B, B-C, A-B-C, and A-B-C-D relations, respectively. The test blocks were always presented in the same order (A-B, B-C, A-B-C, and A-B-C-D).

Testing A-B relations. The first block of eight test trials to which participants were exposed contained the A-B relations. For research purposes, the coin on the left was always referred to as A and the coin on the right referred to as B. At the beginning of the first trial in each session, the child was provided with the following generic instruction: "We are going to play a birthday game. I want you to imagine that it is your birthday today, and you have to go to the shops to get sweets for your birthday party." (This instruction was omitted on all subsequent trials.) Participants were then informed that the researcher could not tell them if their answers were correct or incorrect. All trials were preceded by a question that specified the trial type; for example, "If this coin (researcher pointed to A) buys more sweets than this coin (researcher pointed to B), which would you take to buy as many sweets as possible?"

In each block of A-B trials, there were two exposures to each of four trial types presented in a quasi-random order. These trial types may be described as follows: A buys less (sweets) than B; B buys less than A; A buys more than B; and B buys more than A (see Table 1). Two of the A-B trial types involved the researcher pointing to the A coin first (e.g., A buys more than B) and going from left to right, whereas the remaining two involved the researcher pointing to B first (e.g., B buys more than A) and going from right to left. Irrespective of performance, all participants proceeded immediately to the B-C test trials.

Testing B-C relations. The B-C trial types and the manner in which they were tested were identical to the A-B relations (see Table 1), and the same set of coins was employed. The only difference in the testing of these two repertoires was that the previously used coin B and a novel coin C were presented. During these trials, coin B was always placed to the left of coin C. Once again, irrespective of performance, all participants proceeded immediately to the A-B-C test trials.

Testing A-B-C relations. Testing the A-B-C trial types involved three coins placed on the table in the sequence A then B then C. This testing involved the same three coins that had been employed in the testing of the two previous relational repertoires. An example of a question that comprised an A-B-C trial type is “If this coin (researcher pointed to A) buys less sweets than this coin (researcher pointed to B), and if this coin (researcher pointed again to B) buys less sweets than this coin (researcher pointed to C), which would you take to buy as many sweets as possible?” The four A-B-C trial types were as follows: A buys less than B, which buys less than C; C buys less than B, which buys less than A; A buys more than B, which buys more than C; and C buys more than B, which buys more than A (see Table 1). Within each block of eight test trials, therefore, the researcher pointed to coin A first on four trials and coin C first on the remaining four trials. Once again, irrespective of performance, all participants proceeded immediately to the A-B-C-D test trials.

Testing A-B-C-D relations. Testing A-B-C-D trial types involved four coins placed on the table in the sequence A then B then C then D. This testing involved the same three coins that had been employed in the testing of the three previous relational repertoires, plus an additional fourth coin (D). An example of an A-B-C-D question is “If this coin (researcher pointed to A) buys more sweets than this coin (researcher pointed to B), and if this coin (researcher pointed again to B) buys more sweets than this coin (researcher pointed to C), and if this coin (researcher pointed again to C) buys more sweets than this coin (researcher pointed to D), which would you take to buy as many sweets as possible?” The researcher pointed first to coin A four times and first to coin D four times. The four A-B-C-D trial types are shown in Table 1.

In order to pass a complete baseline test, each participant was required to produce 28 correct responses out of 32 (i.e., 87.5%) with no more than 1 incorrect response on any one of the four blocks. Three of the children (Participant 5/typically developing and Participants 6 and 7/children with autism) were exposed to several baseline tests. The children who were to be exposed to more than one baseline test were selected before the commencement of the study; thus their initial test performances did not dictate the subsequent number of baseline tests to which they would be exposed. All participants who failed the baseline test(s) proceeded immediately to explicit training of the A-B relations.

Training A-B relations. Training of A-B relations was identical to A-B testing and always commenced with the same set of coins. In the case of children exposed to more than one baseline test, training involved the first set of coins presented in the first test. The only difference between training and testing was that feedback was now provided after each response. The instructions provided at the beginning of each training block were identical to those provided during testing, except that the child was also informed as follows: "This time I can tell you if you are right or wrong." All participants continued with A-B training until they had reached a mastery criterion of eight consecutively correct responses. They then proceeded immediately to explicit training of the B-C relations.

Training B-C relations. Training of B-C relations was identical in format to B-C testing, and the same set of coins was employed. Again, the only difference was that feedback was now provided after each response. Participants continued with the training until they emitted eight consecutively correct responses and then proceeded immediately to explicit training of the A-B-C relations.

Training A-B-C relations. Training of A-B-C relations was identical to A-B-C testing (and used the same coins), except with feedback. The training criterion was the same as before and participants then proceeded immediately to explicit training of the A-B-C-D relations.

Training A-B-C-D relations. Training of A-B-C-D relations was identical to A-B-C-D testing (and used the same coins), except with feedback. The training criterion was the same. After A-B-C-D training, all participants were reexposed to a full baseline test (i.e., all four relational repertoires) with a new set of coins. Completion of this second test phase marked the end of each child's participation in Study 1.

Interobserver Reliability

Independent raters, who were staff members in the children's schools, were present for approximately 30% of the training and testing trials conducted with all of the children. Agreement between the raters and the researcher was reached on all but three training trials and two testing trials and yielded an overall interrater reliability score of 99%.

Results

The following results section is divided according to the two groups of children who participated in the study, those who were typically developing and those diagnosed with autism.

Typically Developing Children (Participants 1, 2, 3, 4, and 5)

Participants 1, 2, 3, and 4 followed the procedure for Study 1 as outlined above and required no additional features of training or testing. Because of the significant overlap among the performances of these 4 children, only the results for Participant 1 are described in full. However, the data obtained with all 4 children are outlined in Figures 1, 2, 3, and 4. Due to procedural differences in her involvement in the study, the results of Participant 5 will be described separately (see Figure 5).

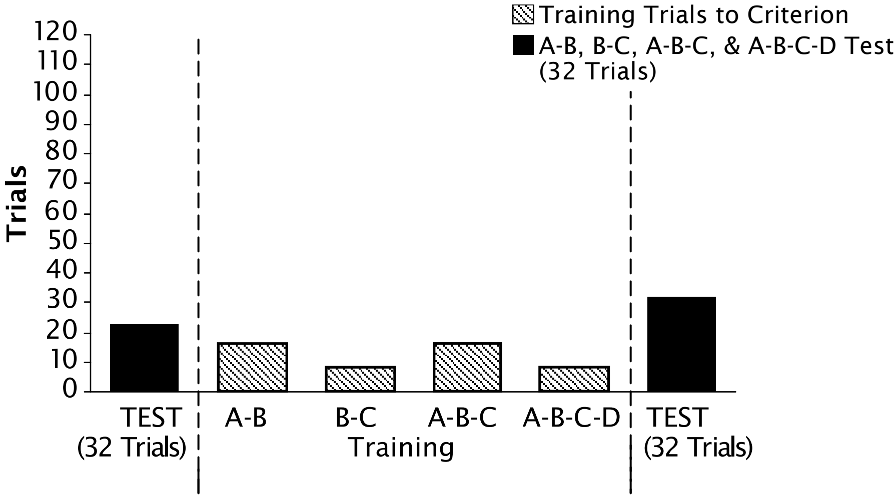

Figure 1. Results for Participant 1 in Study 1.

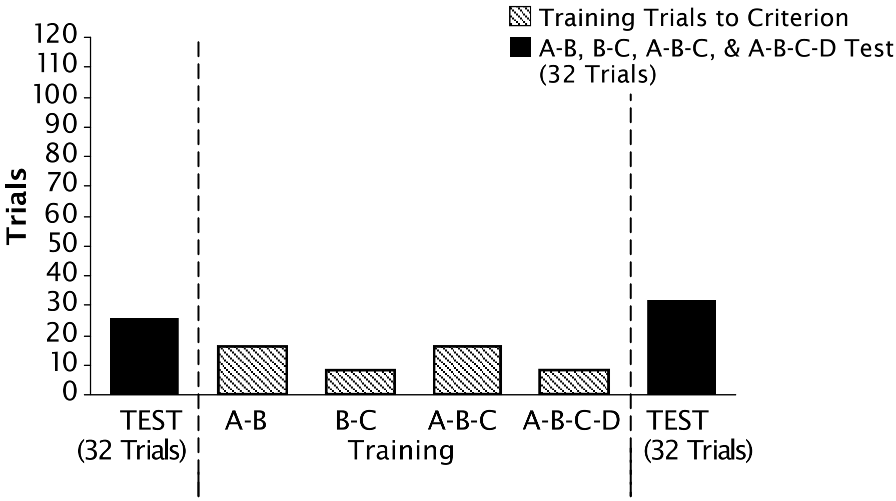

Figure 2. Results for Participant 2 in Study 1.

Participant 1. Participant 1 failed the baseline test with only 23 correct responses and errors across all four relational repertoires. He began explicit training in Session 2 and required only 16 and 8 trials to reach criterion on the A-B and B-C relations, respectively. He similarly completed the A-B-C and A-B-C-D training in only 16 and 8 trials, respectively. On reexposure to the baseline test with novel stimuli (Set 2), he produced perfect responding. Similar performances were recorded with Participants 2, 3, and 4.

Participant 5. Participant 5 failed four separate baseline tests with different sets, with a highest score of only nine correct responses. When first exposed to explicit training, she reached criterion on the A-B relations in only 16 trials and needed only 8 trials of all remaining relations to complete the four training blocks. Not surprisingly, she produced a perfect performance on reexposure to a baseline test. At this point in the procedure, we decided to

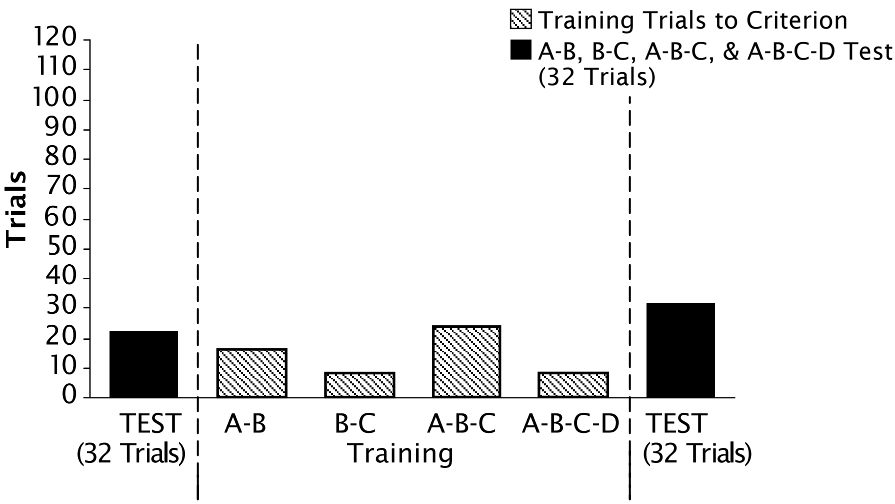

Figure 3. Results for Participant 3 in Study 1.

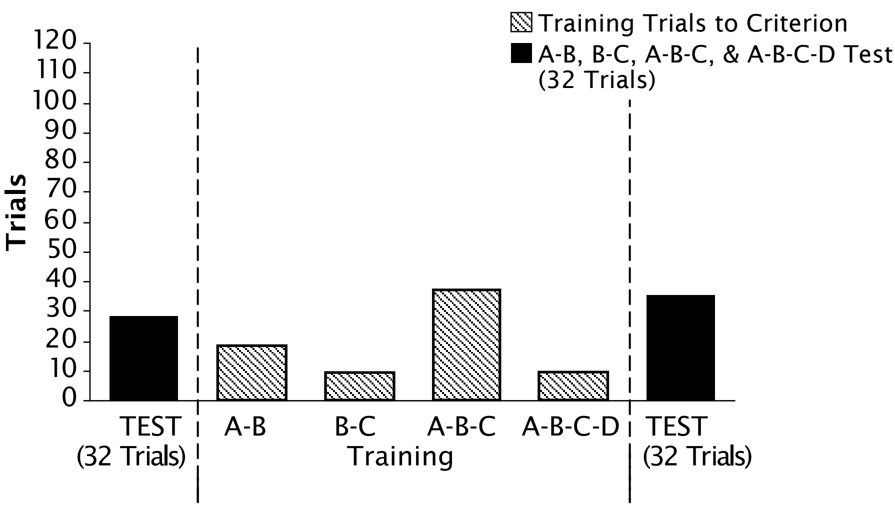

Figure 4. Results for Participant 4 in Study 1.

introduce a fifth coin (identical in size, orange in color, and with a novel pattern) to determine whether the four-coin training would generalize to an increased number of untrained relations (see Table 1). Indeed, she also produced a perfect performance on the novel A-B-C-D-E tested relations. With such a strong set of derived performances recorded thus far, we then decided to test how well these abilities would generalize to changes in the locations of the coins. Specifically, instead of the coins being placed in a linear sequence on the table, they were now placed randomly, and the researcher pointed randomly to each coin in the sequence. All of the baseline relations (i.e., from A-B to A-B-C-D-E) were retested in this way, using a novel set. In this final test, she produced perfect performances on all of the target relations.

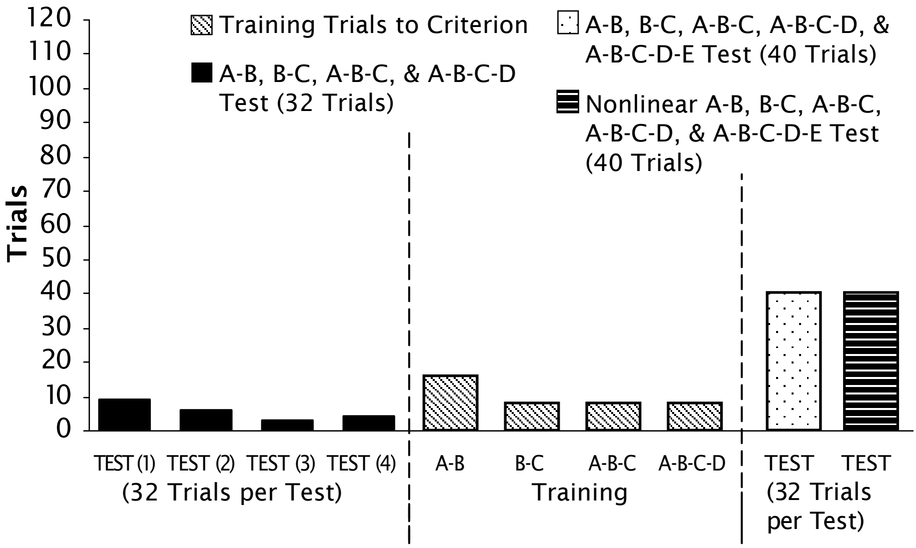

Figure 5. Results for Participant 5 in Study 1.

Children Diagnosed With Autism (Participants 6, 7, and 8)

Two of the 3 children with autism (Participants 6 and 7; see Figures 6 and 7) produced similar performances to the typically developing children and required no additional features of training or testing. The results recorded with Participant 8 (see Figure 8) differed considerably from all of the other children and thus are described in detail.

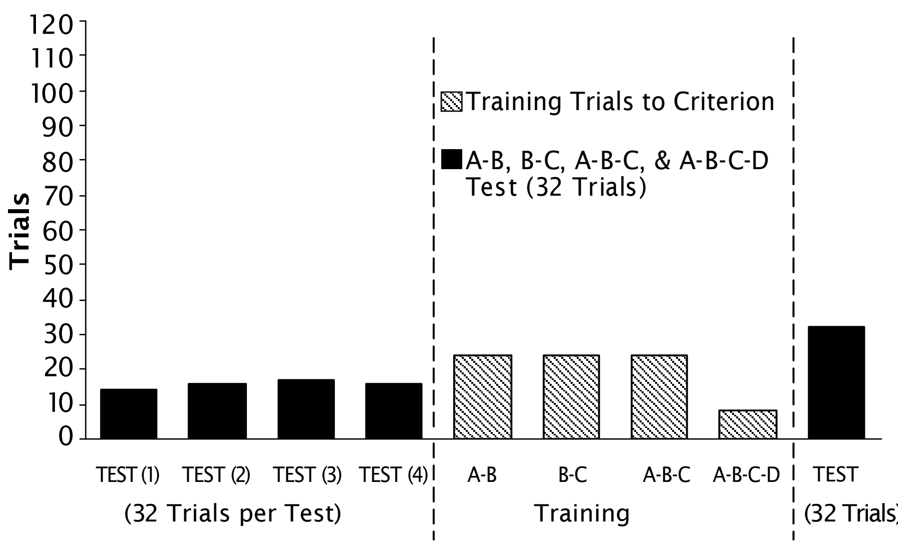

Figure 6. Results for Participant 6 in Study 1.

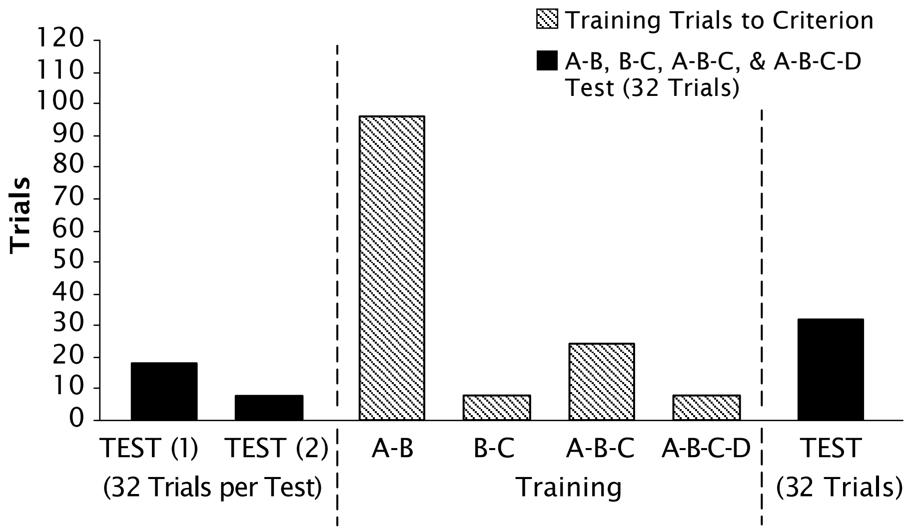

Figure 7. Results for Participant 7 in Study 1.

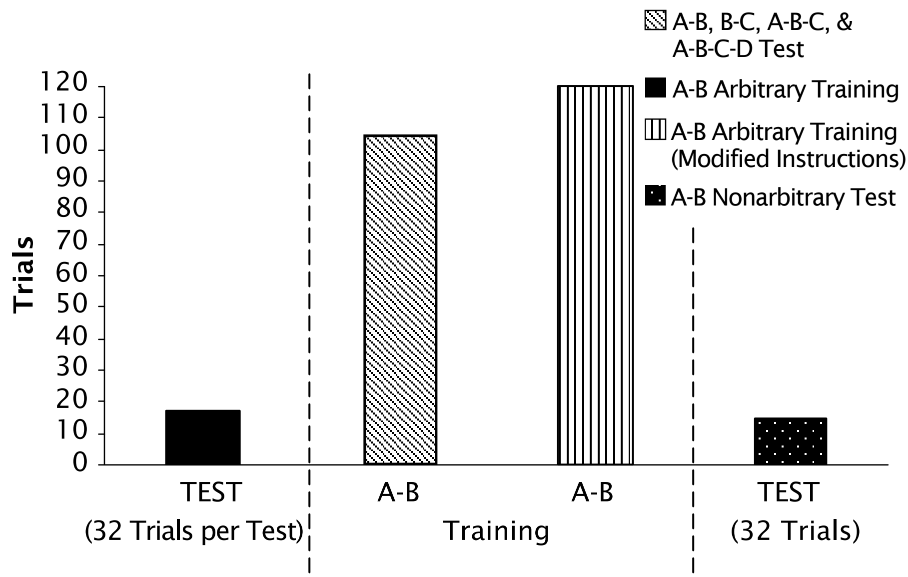

Figure 8. Results for Participant 8 in Study 1.

Participant 8 failed the baseline test and thereafter continued to fail to acquire the A-B relations even after 104 training trials. At this point in the procedure, we decided to introduce a modification to the instruction presented during each trial, because the child did not appear to be attending to the instructions previously provided. During these trials, another reference to each coin was incorporated. Specifically, after the researcher said, for example, "This coin (A) is more than this coin (B)," she then pointed to the first coin (A) and asked the child, "What is this coin?" The correct response in this case involved stating the correct value of the coin (e.g., saying “more”). If the participant responded incorrectly to this question, the response was simply recorded as incorrect. If the participant responded correctly, the researcher then pointed to the second coin (e.g., B) and asked, “What is this coin?” If the participant responded incorrectly to this question, the response was recorded as incorrect. If the child responded correctly to the second question, the researcher asked, “Now, which would you choose?” However, after an additional 120 trials, Participant 8 continued to fail to reach the mastery criterion on the A-B relations.

At this point in the procedure, we decided because of time constraints to discontinue interventions to establish the target A-B relations with this child. However, it was possible to expose the child to one additional test in which his ability to respond to the *nonarbitrary* relations of more than and less than was assessed. Our decision to do this was based on the possibility that his inability to learn the arbitrary A-B relations resulted from deficiencies in responding in accordance with nonarbitrary more-than and less-than relations.

The nonarbitrary test involved the same A-B relations presented previously, except that a number of sweets (e.g., seven) were placed on top of the coin specified in the trial type as *more* and a lesser number of sweets (e.g., two) were placed on top of the coin specified as *less*. The child was then asked, “Which has more?” and “Which has less?” Participant 8 was exposed to four blocks of eight nonarbitrary A-B test trials but produced a total of only 15 out of 32 correct responses. This weak test performance indicated that repertoires of nonarbitrary more-than and less-than relations could not yet be performed by this child in this context and thus offered one possible reason why he repeatedly had failed to reach criterion during explicit training of the arbitrary relations. At this point, the participation of Participant 8 was terminated.

Discussion

Study 1 involved the participation of 5 typically developing children and 3 with autism. All were exposed to the same basic problem-solving task in which responding in accordance with linear, arbitrary more-than and less-than relations with two, three, and four coins was targeted for establishment. Nonlinear coin sequences involving five coins were also presented to 1 typically developing child (Participant 5). All 8 children failed all baseline tests to which they were exposed (some failed as many as four), involving the target A-B, B-C, A-B-C, and A-B-C-D relations. Seven of these children then proceeded very quickly through explicit training of the relations and passed a subsequent test of all relations involving a novel stimulus set. After several hundred A-B training trials and modified instructions, 1 child (Participant 8) continued to fail to reach criterion. However, a very weak performance in a subsequent test of the nonarbitrary A-B relations offered a sound explanation for why the arbitrary relational training had not succeeded. With the exception of Participant 8, the results overall showed few differences between the typically developing children and those with autism in establishing and generalizing the arbitrary more-than and less-than relations. The current findings replicated and extended those reported previously by Barnes-Holmes, Barnes-Holmes, Smeets, Strand, et al. (2004) and Berens and Hayes (2007), with the inclusion of children with autism.

Study 2

One issue raised by the results from Study 1 and referred to previously is the possibility that because the same trial types were employed in both training and testing, the target performances were not truly derived. In other words, most of the children were explicitly trained on A-B, B-C, A-B-C, and A-B-C-D relations, and all of the test trials contained only these relations. Although the derived performances recorded by Participant 5 with five coins contradicts this possibility, stronger evidence for derivation would emerge if one could examine relations that had not been included in previous testing or training. Specifically, although it had not been targeted explicitly, a four-coin sequence contains a pair of transitive relations between the nonadjacent B and D coins. Thus, testing these relations directly provided an opportunity to investigate novel derived transitive relations that had not been explicitly trained. Although the study of transitive relations (referred to as *transitive inference*) has attracted little or no interest from behavioral psychologists, it has been a primary concern of cognitive researchers for many years. Thus, this section briefly reviews the cognitive literature on acquisition of transitive inference in children.

Transitive inference (i.e., the derivation of transitive relations) has been studied extensively by cognitive researchers in the context of the *three-term series problem* (Hayes-Roth & Hayes-Roth, 1975). Successfully deriving transitive relations is associated with the ability to observe hierarchical relationships among stimuli (Moses, Villate, & Ryan, 2006). Several studies have investigated this phenomenon with young children. For example, Russell, McCormack, Robinson, and Lillis (1996) distinguished between sequential (i.e., linear) and nonsequential (i.e., nonlinear) transitive inference tasks. Sequential transitive inference tasks are employed most commonly with children and involve the presentation of item sequences similar to those employed in Study 1. According to Russell et al., this type of task is easier to solve than nonsequential tasks because when the pairs are trained in sequence the transitive pairs also can be determined using the same sequence. For example, if a participant is trained that A > B, B > C, C > D, and D > E, then it logically follows in the linear sequence that B > D and C > E. The results reported by Russell et al. indicated that children as young as 5 years of age can respond in accordance with transitive relations when presented in sequential transitive inference tasks. Indeed, similar work conducted by Bryant and Trabasso (1971) suggested that transitive inference can be demonstrated by children as young as 4 years old.

Other cognitive researchers, however, have argued that young children are *not* capable of transitive inference, because these skills do not develop until a later age. According to Piaget, for example, the ability to make transitive inferences reflects the use of special mental operations that develop only as children of 7 or 8 years of age enter the concrete operational stage (Mims, Cantor, & Riley, 1983; Piaget, Inhelder, & Szeminska, 1960). However, Bryant and Trabasso (1971) argued that a critical weakness of Piaget's work concerned the absence of a test to determine that the children could in fact remember the relevant information in the premises and derive the other relations (e.g., the symmetrical relations) necessary for determining the more complex transitive relations. Hence, it is difficult to determine the precise level of special mental operations children need in order to demonstrate transitive inference or the normal age at which these emerge. Indeed, there appear to be no published studies of attempts to investigate transitive inference in populations who present with cognitive impairments.

Study 2 was also an attempt to build upon the repertoires that had been established previously for the children, by extending the target relations to five coins. The basic preparation employed in Study 2 was identical to Study 1 except that the transitive relations among the nonadjacent coins denoted as B and D were targeted and five-coin sequences also were presented.

Method

Participants

Seven of the 8 children involved in Study 1 also participated in Study 2. Participant 8 did not participate in Study 2.

Setting and Materials

The setting and materials employed in Study 2 were identical to Study 1 but involved 6 additional coins that constituted a 5th coin in each of the six stimulus sets. Hence, a total of 30 coins were now employed. All 6 new coins (always referred to as coin E) were orange in color, and each was distinguished by a particular pattern.

Programmed Consequences

The programmed consequences employed in Study 2 were identical to those outlined for Study 1.

Procedure

Study 2 consisted of testing and training of the B-D relations within four- and five-coin sequences and was conducted in the next session following the completion of Study 1. The children were presented with a four-coin A-B-C-D sequence as in Study 1 but were now required to derive the transitive B-D relations contained therein. In other words, the target derivation of B-D relations within a four-coin sequence was a type of generalization of the existing relations that had been established from the previous training and testing. Put simply, if accurate responding to A-B-C-D relations was present, could the children also derive the composite B-D relations within this sequence?

In addition, the children in Study 2 were presented with a five-coin A-B-C-D-E sequence (presented only to Participant 5 in Study 1) and were required to derive the transitive B-D relations there. Again, this modification was designed to determine whether the children's competence on four-coin sequences would generalize to five-coin sequences and whether competence on the latter also involved accuracy on the derivation of the composite transitive B-D relations. All of the trials presented during the training and testing that comprised Study 2 involved B-D relations in four- or five-coin sequences.

Testing B-D relations with four coins. Each child was first exposed to a baseline test of the more-than and less-than B-D relations within a sequence of four coins that was identical to the A > B > C > D trial types from Study 1.

However, instead of asking, “Which would you take to buy as many sweets as possible?”, the researcher asked, “Which would you choose—this coin (researcher pointed to coin B) or this coin (researcher pointed to coin D)—to buy as many sweets as possible?” Hence, the transitive baseline test consisted of four B-D trial types (two more-than and two less-than trials). Although the researcher may have pointed to either the A coin or the D coin first in presenting the trial type (e.g., in the trial A < B < C < D, she pointed to A first), she always pointed to the B coin first when asking the participant to determine the transitive relations between the B and D coins (i.e., the child was always asked to choose between B and D, not vice versa). This procedure ensured that the participant could not determine the transitive relations by simply using the sequence of pointing presented by the researcher.

A correct transitive response was recorded when a child pointed to the correct coin or handed it to the researcher. For example, during the trial A < B < C < D, the correct response involved selecting the D coin (because D was worth more than B). If the participant identified two or more coins, this response was deemed incorrect. In order to pass the baseline test, each participant was required to produce seven correct responses (out of eight) without producing more than one incorrect response on any one B-D trial type. Similar to Study 1, some of the children were again exposed to several baseline tests. All participants, irrespective of performance, proceeded immediately to a further baseline test of the B-D relations among five coins.

Testing B-D relations with five coins. The baseline test of the B-D relations among five coins was identical to the previous test except that five coins (denoted as A-B-C-D-E) were presented on all eight trials (see Table 1). The children who had received multiple exposures to the baseline test involving four coins were exposed to the same number of baseline tests involving five coins.

Training B-D relations with four coins. Children who failed the baseline transitive test with four coins were exposed to explicit training of the B-D relations in this sequence. Once again, this training was identical to testing, except that the trials were consequated with corrective feedback. All B-D training was conducted in blocks of eight trials and continued until each child reached the mastery criterion of eight consecutively correct responses. Participants then proceeded immediately to another test of the B-D relations among four coins involving a new stimulus set. Those who failed this test were retrained and tested until accurate performances were recorded on the B-D relations on a novel set of four coins and in the absence of feedback.

Training B-D relations with five coins. Participants who failed the B-D test with five coins were exposed to explicit training of same. Again, this training was identical to the testing, except that corrective feedback was provided on every trial. Participants were required to reach a mastery criterion of eight consecutively correct responses before reexposure to the same test with a novel set.

Interobserver Reliability

The same independent raters from Study 1 were employed in Study 2 to rate approximately 30% of the trials conducted with all children. In Study 2, the raters were present during 270 of a total of 856 training and test trials. Agreement between the raters and the researcher was reached on all but 5 training trials and 1 test trial. The interrater reliability score recorded for this study was 97.7%.

Results

Typically Developing Children

The data obtained with Participants 1, 2, 3, 4, and 5 in Study 2 are presented in Figures 9 through 13, respectively. Once again, the results recorded with Participant 2 are described in full, whereas those of Participants 1, 3, and 4 are described only in terms of the aspects of their performances that differentiated them from Participant 2. Again, Participant 5's results are also described due to procedural differences in her participation in the study.

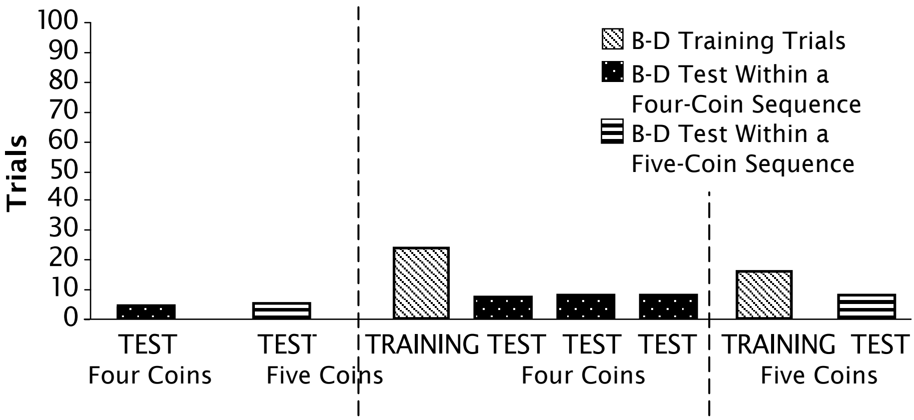

Figure 9. Results for Participant 1 in Study 2.

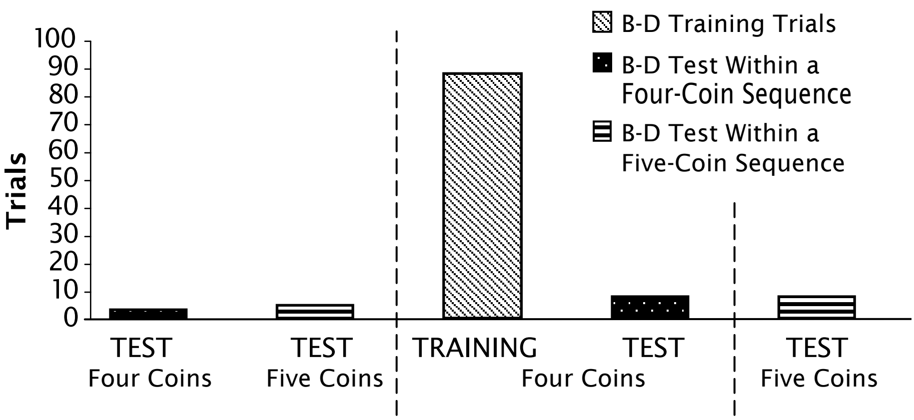

Figure 10. Results for Participant 2 in Study 2.

Participant 2. Participant 2 failed each of the four- and five-coin B-D tests with three of eight and five of eight correct responses, respectively. He required 88 trials to reach criterion on the B-D relations with four coins and passed a subsequent test with novel coins with a perfect performance. When reexposed to a test of the B-D relations among five coins with a new set, he produced a perfect performance. Participant 3 required two separate training exposures to the B-D relations with four coins, and Participant 1 required some explicit training of the B-D relations among five coins.

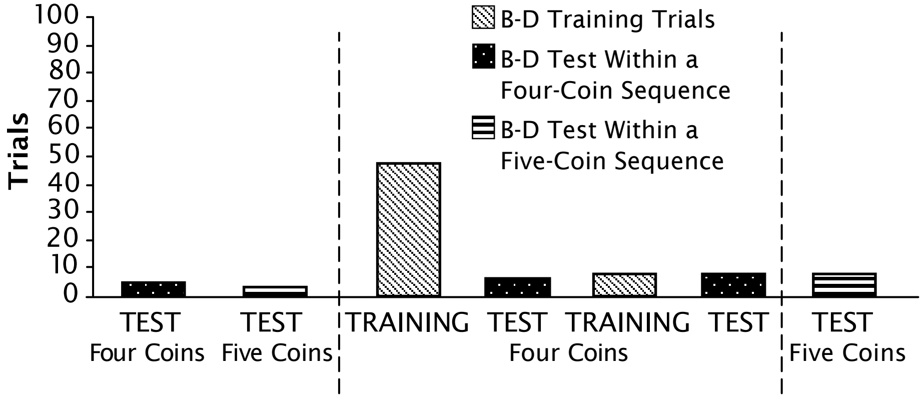

Figure 11. Results for Participant 3 in Study 2.

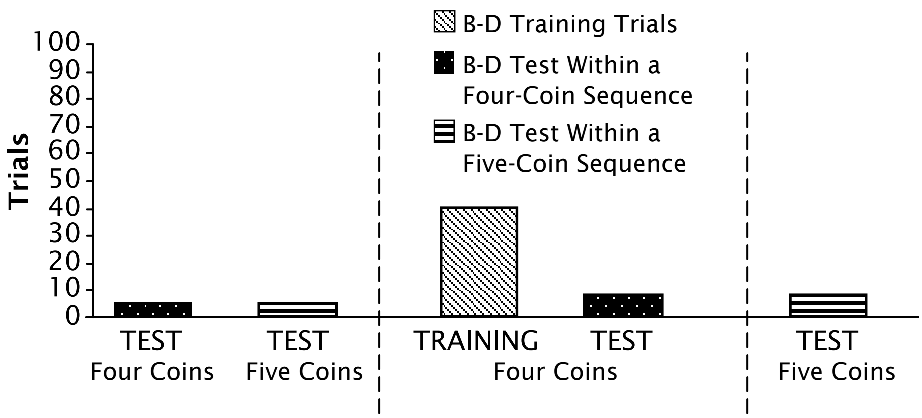

Figure 12. Results for Participant 4 in Study 2.

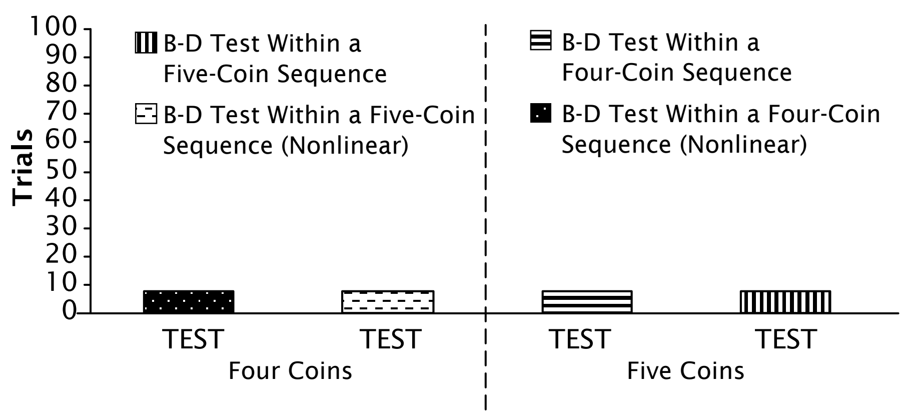

Figure 13. Results for Participant 5 in Study 2.

Participant 5. Participant 5 produced a perfect performance on the B-D test with four coins. Based on her performances in Study 1, we decided at this point to expose her to a nonlinear test of the same relations with a new set of four coins. Again, she produced a perfect performance. The child was then tested on the B-D relations with five coins and again produced a perfect performance. Finally, she was exposed to a nonlinear test of the same relations with a new set of five coins and produced a perfect performance.

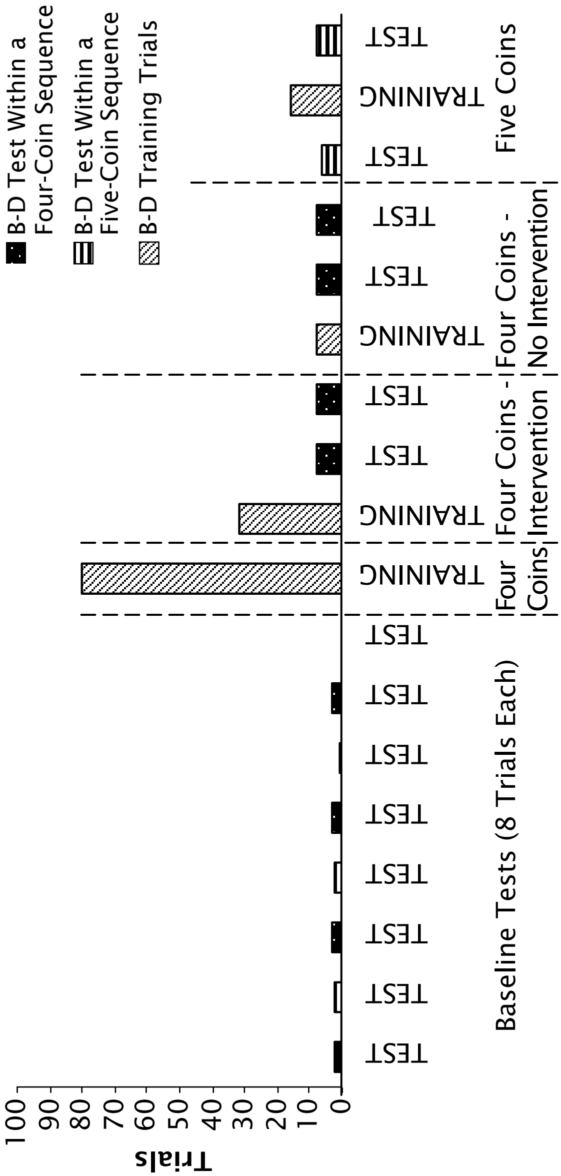

Figure 14. Results for Participant 6 in Study 2.

Children Diagnosed With Autism

The data obtained with Participants 6 and 7 are outlined in Figures 14 and 15, respectively. Both of these children showed greater difficulty in the establishment of the B-D relations with four coins during training than did their typically developing counterparts, and an additional intervention to establish these relations was implemented. The details of this intervention are provided in the context of the data for Participant 6.

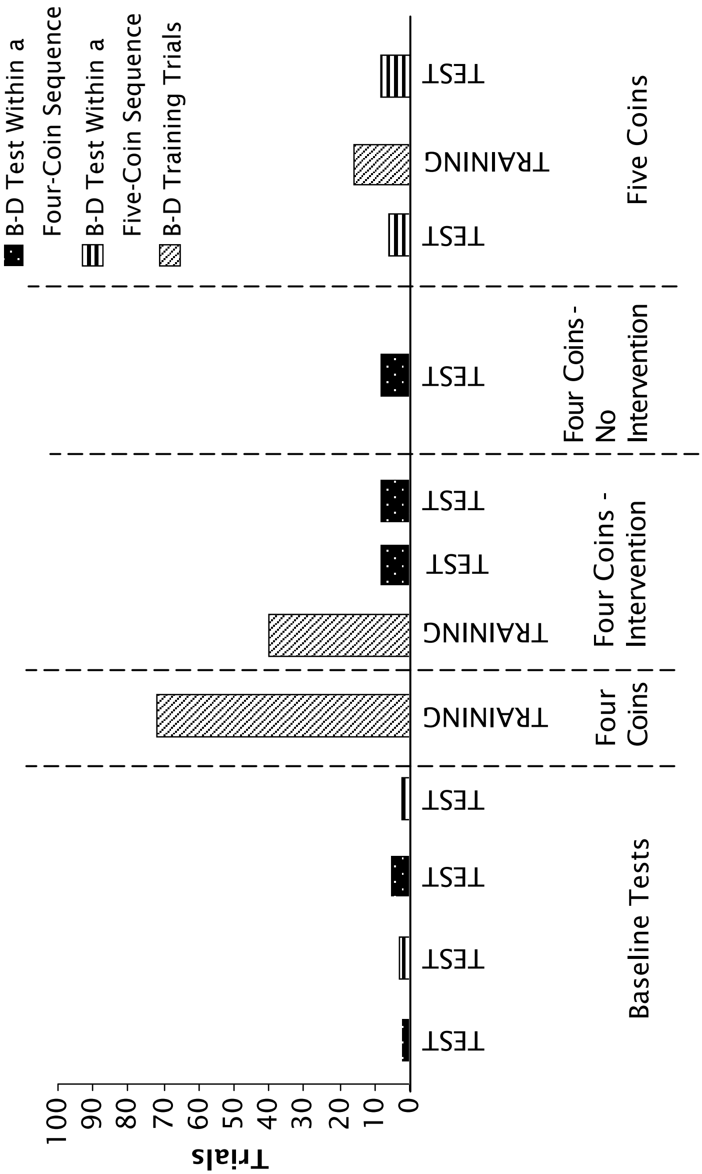

Figure 15. Results for Participant 7 in Study 2.

Participant 6 failed four baseline tests of the B-D relations within four- and five-coin sequences. After 80 training trials across eight sessions, he still had not reached the mastery criterion on four coins. The primary difficulty demonstrated by the child appeared to be a lack of attention to the stimuli (a problem also noted by Bryant & Trabasso, 1971). In order to remediate this difficulty, an intervention was devised in which the child's attention was drawn explicitly to the B and D stimuli. This consisted of placing the four coins on a plain white card and drawing two circles, one around the B coin and one around the D coin. After 32 trials involving this intervention, Participant 6 reached criterion on the B-D relations with four coins. He was then immediately tested using the same set (and the intervention) and produced a perfect performance. On a subsequent test involving a novel set, again with the intervention, he also produced a perfect performance. In Session 17, he was exposed to explicit training of the B-D relations with four coins without the intervention (i.e., removing the card so that each of the four coins were simply placed on the table from A to D as before) and reached criterion in only 8 trials. When thereafter reexposed to the test (without the intervention), he produced perfect performances with trained and novel stimuli. Participant 6 subsequently was tested on the B-D relations with five coins. After failing the test, he was explicitly trained on the target relations and reached criterion in only 16 trials. He then was tested with a new set of five coins and produced a perfect performance.

The intervention involving the cards employed with Participant 6 was also necessary with Participant 7, who had failed to reach criterion on the B-D relations with four coins after 72 trials. With this child, the intervention appeared to be equally effective.

Although Study 1 demonstrated generalized responding in accordance with arbitrary more-than and less-than relations with four coins, 6 of the children in Study 2 (typically developing and those with autism) failed the baseline tests of the B-D relations with four (and five) coins. Successful test performances of these transitive relations on four novel coins were demonstrated after relatively limited explicit training with four of the children (all typically developing). Three of these children (Participants 2, 3, and 4) then also passed the five-coin transitive test immediately with novel coins, with the remaining child (Participant 1) requiring some explicit training on these relations. The 2 children with autism (Participants 6 and 7) required an additional intervention in order to complete training on the four-coin sequence and also required some explicit training of the relations within a five-coin sequence. Participant 5 was the only child who had been exposed to a five-coin sequence in Study 1, and she immediately demonstrated the B-D relations with both four- and five-coin sequences in Study 2. This child also demonstrated generalized responding on the B-D relations in nonlinear sequences. The results of Study 2, therefore, demonstrate the expansion of the comparative repertoires established previously in terms of both number of coins and responding in accordance with the transitive B-D relations contained therein.

General Discussion

Eight children, 5 typically developing and 3 diagnosed with ASD, participated in this study. In Study 1, all children failed all baseline tests involving the arbitrary more-than and less-than relations among four coins presented in a linear sequence to which they were exposed. Some of the children were exposed to multiple baseline tests. This indicates that the target performances did not emerge as a result of repeated baseline testing alone. After explicit training of the composite relational performances, 7 of the children demonstrated the target performances on a novel set of four coins (Participant 8, who had autism, did not). Overall, few differences were observed between the training and testing profiles of the typically developing children and those with autism, apart from the need for more extensive training of the A-B relations with the latter group. Furthermore, no additional interventions were required with any of the children to facilitate the training of the target relations. The results of Study 1 overall indicated the successful establishment of the target arbitrary relations of more than and less than in a linear sequence using multiple exemplar training with these two samples of young children. These results are similar to those reported in the original study by Barnes-Holmes, Barnes-Holmes, Smeets, Strand, et al. (2004) and further support the utility of exemplar training in the context of arbitrary relations. The findings also support the RFT assumption that derived relational responding is generalized operant behavior.

The 7 children who demonstrated successful test performances in Study 1 also participated in Study 2. Although all had previously passed tests of the more-than and less-than relations among four coins, 6 of the children failed to demonstrate derivation of the transitive B-D relations within four- and five-coin sequences in the baseline tests of the latter study. Once again, the use of multiple baseline tests with several children indicated that the target transitive performances were not established by repeated testing alone. Although 4 of the typically developing children trained quickly on the four-coin B-D relations, the 2 children with autism required an intervention that emphasized the B and D coins. After this training, all 6 children readily passed a four-coin test of B-D relations on a novel set. Three of the typically developing children, but none of those with autism, then passed the B-D relations within five coins without explicit training. Although the 3 remaining children did not pass the five-coin B-D test immediately, only very limited explicit training of these relations was required.

Participant 5 proceeded more quickly through the training of the target relations in Study 1 than the other children, in spite of four failures on the baseline tests. After passing the four-coin test on a novel set, Participant 5 generalized immediately to five coins and to the presentation of the coin sequences in a random (nonlinear) manner. These performances appeared to support perfect responding (in the absence of training) on linear and randomized tests of the B-D relations with four and five coins in Study 2.

The overall performances of the typically developing children and those with autism (with the exception of Participant 8) did not differ considerably across both studies reported here. The two notable areas in which the children with autism appeared to produce weaker performances can be summarized as follows: (a) In Study 1 they required more extensive training of the A-B relations, and (b) in Study 2 it was necessary to highlight the B and D coins during explicit training of the B-D transitive relations in a four-coin sequence. These differences appeared to be readily rectified with the exemplar training and the brief intervention to highlight the B and D coins, in order to produce similar performance outcomes across the two groups of children.

The specific intervention employed to facilitate the transitive B-D relations with four coins for the children with autism in Study 2 was based on a suggestion made by Bryant and Trabasso (1971). In their original work, the researchers reported that “lower performance on the BD test may be attributed to a failure in memory rather than to inferential difficulty” (p. 458). In the current study, however, the emphasizing of the B and D coins was conducted simply to draw the child’s attention to the specific coins. It is difficult then to determine whether this intervention permitted the circling of the coins to function as discriminative stimuli for attending to those coins directly, or whether it allowed the coins to function as contextual cues for the subsequent relational response. In either case, it is not necessary to employ the concepts of *memory failure* or *inferential difficulty,* as suggested by Bryant and Trabasso, in order to account for the success of this intervention.

The performances of Participant 8 differentiated him from both the typically developing children and the other children with autism. Although he was exposed to extensive training of the A-B relations as well as simplified instructions, he continued to fail to reach criterion. Although this failure precluded him from further participation in the study, he was then exposed to a test of the nonarbitrary more-than and less-than relations. According to RFT, failure to pass this test would render it unlikely that he would readily learn arbitrary more-than and less-than relations (Barnes-Holmes, Barnes-Holmes, & Cullinan, 2001). Two features of the current data support this suggestion, albeit indirectly. First, it seems reasonable to assume that the children who demonstrated arbitrary relations would have easily passed tests of nonarbitrary relations, had they been presented. Second, the only child who failed to derive arbitrary relations also failed to demonstrate the nonarbitrary relations. Therefore, the failure of the latter child suggested that he had failed to learn the arbitrary relations at least in part because the nonarbitrary relations were not yet fully established. These results also support other positive outcomes that have been obtained when nonarbitrary interventions have been employed as a means of establishing arbitrary relations (Barnes-Holmes, Barnes-Holmes, Smeets, Strand, et al., 2004; Berens & Hayes, 2007). Unfortunately, Participant 8 became unavailable for further participation in the current study, and thus we were unable to continue to investigate his repertoires of nonarbitrary more-than and less-than responding.

Barnes-Holmes, Barnes-Holmes, Smeets, Strand, et al. (2004) contended that one possible criticism of their study was that, because the children were trained and tested on the same four trial types using only 2 or 3 coins, one might argue that the test performances, even on novel stimulus sets, were not indicative of *truly derived* relational responding. Barnes-Holmes, Barnes-Holmes, and Smeets (2004) suggested that two ways to address this issue involve (a) increasing the number of coins and (b) modifying the sequences in which the coins are presented (i.e., between linear and random). In the Barnes-Holmes, Barnes-Holmes, and Smeets study, for example, the children required many training exemplars to derive the opposite relations between 2, 3, 4, and 5 coins but few or no exemplars of training for 6, 7, 8, 9, or 10 coins. Both of these strategies were employed to some degree in the current studies. The high levels of generalization reported for the majority of children here (especially for Participant 5) are consistent with the findings reported in the

Barnes-Holmes, Barnes-Holmes, and Smeets study. However, one might still argue that this does not circumvent the criticism because the target relations in all cases remained the same during training and testing.

The focus on transitive relations in Study 2 was designed explicitly to address this issue because transitive relations between a nonadjacent pair were not targeted in Study 1 but were tested immediately at the beginning of Study 2. The results from the latter study, however, indicated that the majority of children did not pass the baseline B-D tests without explicit training. One might still argue, therefore, that the children's performances for the most part were not truly derived. Participant 5, however, remains the exception. She did in fact pass the B-D tests immediately and without explicit training, suggesting that at least in her case, the latter performances were derived and probably based in part on her training and testing in the previous study. Additional research, however, will be necessary to investigate the degree of derivation required to pass the types of arbitrary tests presented here. Nonetheless, the current studies provide further evidence of the facilitation of strong generalized performances with typically developing children and those with autism.

A related criticism of the current research concerns the possible influence of the children's preexperimental histories relative to their explicit experimental learning experiences on the emergence of the current performances. Because the children in the current studies ranged in age from 6 to 10 years old, and the children with autism in particular had each had several years of experience in an ABA facility, it is reasonable to assume that at least some of the children already had extensive relational repertoires that may have overlapped with those targeted here. Hence, one may only argue that the deficits currently observed were specific to some extent to the present research context. However, it remains the case that all of the children produced very weak baseline performances that were quickly ameliorated with the exemplar training provided. Further research, perhaps with younger children or those who are more profoundly developmentally delayed, will be necessary to determine whether the same interventions would be equally effective in the context of establishing relational repertoires ab initio. Nonetheless, the current outcomes remain a valuable contribution to research on the broad application of behavioral principles and the specific utility of work on derived stimulus relations.

The issues of age and necessary prerequisite skills for deriving transitive relations naturally emerge from the current work. All of the children who participated in Studies 1 and 2 were between 6 and 10 years old. Indeed, the only child who failed to proceed from Study 1 to Study 2 was the oldest child of all, had autism, and was within the lowest verbal category according to the CABAS system. While it is difficult to determine the precise relationship between age and verbal abilities, the data here appear to better support Piaget's belief that children under 7 years of age probably cannot do transitive inference. However, the speed and ease with which some of the 6-year-olds in this study learned the target relations does suggest that similar interventions might work equally well with younger children. For example, if nonarbitrary relations were targeted first, there would seem to be no reason to believe that success on arbitrary relations would not follow, even with younger children. This then would also appear to support the suggestions made by Bryant and Trabasso (1971) that children as young as 4 years of age may demonstrate transitive inference. Indeed, the success we had here with children with autism does support this conclusion. For RFT, therefore, much can be gained when interventions are adapted appropriately to an individual's level of verbal ability and when the interventions best target core relational deficits.

Although the current work was generated directly by RFT, alternative interpretations of the findings are also possible, and similar arguments have been addressed by the current authors in related research (Barnes-Holmes, Barnes-Holmes, Smeets, Strand, et al., 2004). For example, it could be argued that on the initial relations (excluding the transitive relations) the children simply learned to respond to the first coin when the researcher said the word *more* and to the last coin when presented with the word *less.* In this case, therefore, one could argue that the training simply established two stimulus classes, whose S+ and S- functions were determined by the spoken words *more* and *less*. Consistent with this view, one might then cite the work of Vaughan (1988), who established two stimulus classes by means of repeated reversal training with pigeons. Nevertheless, the use of novel stimuli throughout the present research differentiates it considerably from Vaughan's work, because responding to the novel sets came under the contextual control of *more* and *less* in the absence of explicit reinforcement. Furthermore, no such argument could account for the derived transitive performances of Participant 5 in this study, because on these trials, the words *more* or *less* did not provide a ready basis for accurate responding. In conclusion, therefore, even if one chooses not to interpret the findings in RFT terms, the results from both studies significantly extend existing research.

The current work has implications for basic research on the development of language and cognition in young, typically developing children and those with autism, and for the integration of these interests to applied behavior analysis. At the very least, the research speaks to educational programs (in typical or special educational facilities) in which arbitrary comparative relations are established. For example, young children spend long periods at school learning mathematics. This training comprises perhaps many hundreds of examples of more-than and less-than relations (e.g., with addition and subtraction) that consume a large proportion of their time at school. The RFT approach to relational responding, however, suggests that interventions aimed at targeting the relational skills directly and establishing these across exemplars may have the potential to minimize the training resources considerably and also could generate relational repertoires that are both more complex and more flexible. Although many of these ideas have yet to be scrutinized experimentally, the current study offers one more reason why this important applied possibility should be explored.

References

BARNES-HOLMES, Y., BARNES-HOLMES, D., & CULLINAN, V. (2001). Education. In S. C. Hayes, D. Barnes-Holmes, & B. T. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 181–196). New York: Kluwer Academic/Plenum.

BARNES-HOLMES, Y., BARNES-HOLMES, D., & SMEETS, P. M. (2004). Testing and training relational responding in accordance with the relational frame of opposite in young children. *International Journal of Psychology and Psychological Therapy, 4,* 559–586.

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

BERENS, N. B., & 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 inference and memory in young children. *Nature, 232*, 456–485**.**

CARR, D., & FELCE, D. (2000). Application of stimulus equivalence to language intervention for individuals with severe linguistic disabilities. *Journal of Intellectual & Developmental Disability, 25*(3), 181–205.

HAYES, S. C., & WILSON, K. G. (1996). Criticisms of relational frame theory: Implications for a behavior analytic account of derived stimulus relations. *The Psychological Record, 46,* 221–236.

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

HAYES-ROTH, B., & HAYES-ROTH, F. (1975). Plasticity in memorial networks. *Journal of Verbal Learning and Verbal Behavior, 14*, 506–522.

KOEGEL, L. K., CARTER, C. M., & KOEGEL, R. L. (2003). Teaching children with autism self-initiations as a pivotal response. *Topics in Language Disorders, 23*(2), 134–145.

MIMS, M., CANTOR, J. H., & RILEY, C. A. (1983). The development of representation skills in transitive reasoning based on relations of equality and inequality. *Child Development, 54,* 1457–1469.

MOSES, S. N., VILLATE, C., & RYAN, J. D. (2006). An investigation of learning strategy supporting transitive inference performance in humans compared to other species. *Neuropsychologia, 44,* 1370–1387.

PIAGET, J., INHELDER, B., & SZEMINSKA, A. (1960). *The child's conception of geometry.* Oxford, England: Basic Books.

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.

VAUGHAN, W., JR. (1988). Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 36–42.