Paper:

May, R., Stewart, I., Baez, L., Freegard, G. & Dymond, S. (2017). Arbitrarily applicable spatial relational responding. *Journal of the Experimental Analysis of Behavior, 107*(2), 234-257.
http://dx.doi.org/10.1002/jeab.250

Arbitrarily Applicable Spatial Relational Responding

Richard J. May[a], Ian Stewart[b], Luisa Baez[a], Gary Freegard[c], and Simon Dymond[c, d]

[a] School of Psychology, University of South Wales, Pontypridd, CF37 1DL, United Kingdom.

[b] School of Psychology, National University of Ireland, Galway, Ireland.

[c] Department of Psychology, Swansea University, Swansea, SA2 8PP, United Kingdom.

[d] Department of Psychology, Reykjavík University, Menntavegur 1, Nauthólsvík, 101 Reykjavík, Iceland.

Address correspondence to: Richard J. May, School of Psychology, University of South Wales, Pontypridd, CF37 IDL, United Kingdom. Email: Richard.may@southwales.ac.uk

Abstract

Spatial reasoning, in which a person infers novel spatial relationships based on trained relationships, can be conceptualized as arbitrarily applicable spatial relational responding. Here, we conducted two experiments to develop and validate, for the first time, a laboratory procedure to establish arbitrarily applicable spatial relational responding in adult humans. In Experiment 1, participants were trained on nonarbitrary spatial relational tasks designed to establish contextual cues for *left of, right of, above,* and *below*. Contextual cues were then used to train a series of arbitrary spatial relations involving four abstract shapes. Following training in a subset of arbitrary relations (A is *left of* B, B is *above* C, C is *right of* D), subsequent testing examined the emergence of untrained spatial relations (B is *right of* A, C is *below* B, D is *left of* C, D *is below* A and A is *above* D). When absent in initial tests, spatial relational responding was facilitated by a remedial training procedure incorporating nonarbitrary relational guidance. Participants showed patterns of spatial relational responding consistent with test relations. In Experiment 2, a variant reversal design yielded predictable, reversed spatial relational responses. Overall, the present procedures represent the first empirical demonstration of arbitrarily applicable spatial relational responding and thus, arguably, the first functional analytic model of spatial reasoning.

Keywords: spatial relations, relational frame theory, nonarbitrary, arbitrary, reversal, humans.

Spatial reasoning is a domain of human activity that has received considerable attention from cognitive and developmental psychologists. One particularly well studied example is spatial transitivity, a form of *transitive inference* (TI), a phenomenon proposed to be an essential feature of problem solving and reasoning (Vasconcelos, 2008). In a typical spatial TI task, participants initially learn the spatial relationship between particular pairs of stimuli, such as, “A is left of B; B is left of C”. TI is demonstrated when, as a result, they respond correctly or ‘infer’ the spatial relationship between pairs that were not directly encountered together during training, in this case, “A is left of C and C is right of A”. This phenomenon has been studied extensively in the cognitive and developmental psychology literatures where it has been shown that humans readily engage in spatial TI and in accordance with a variety of complex spatial relational patterns (Klauer, Stegmaier, & Meiser, 1997; Krumnack, Bucher, Nejasmic, Nebel, & Knauff, 2011; Nejasmic, Bucher, & Knauff; 2015).

In general, developmentally focused researchers typically track the emergence of spatial TI and other forms of spatial reasoning over time (i.e., Pears & Bryant, 1990) or the relationship they have to intellectual development (i.e., Bryant, 1974), whereas cognitively focused researchers have used observations of such patterns to infer hypothetical mental structures (Halford, Wilson, & Phillips, 2010; Knauff, Strube, Jola, Rauh, & Schlieder, 2004). Such accounts propose that because information about the correct choices is not explicitly available from stimuli in actual physical relationships with one another in the testing context (and often in either the training or testing contexts), participants solving the tasks must rely on mental representations, models, or images of the stimuli in order to make the correct inferences (Gazes, Lazareva, Bergene, & Hampton, 2014). The aims and assumptions of these research paradigms differ from those of behavior analysis however. Behavior analytic researchers emphasize the need for scientific analysis to specify manipulable variables in order to achieve *influence* over observable behavior (Hayes & Brownstein, 1986). In other words, they aim to provide a functional analytic account of the origins and development of behavior, including spatial reasoning.

Within behavior analysis, one way in which spatial reasoning can be interpreted is as a form of contextually controlled relational responding. Contextually controlled or arbitrarily applicable relational responding (AARR) refers to patterns of "relating" (responding to one stimulus in terms of another) that are controlled by relation-specifying contextual cues as opposed to the nonarbitrary properties of the relata involved (e.g., a dime is worth 'more than' a nickel despite being the physically smaller of the two coins; Hayes, Barnes-Holmes, & Roche, 2001; Stewart, 2016). Such cues typically acquire their discriminative properties, initially, in nonarbitrary contexts (see, for example, Berens & Hayes, 2007; Hayes, Fox, Gifford, Wilson, Barnes-Holmes, & Healy, 2001; Stewart & McElwee, 2009). A child may be taught, for example, to choose the bigger of two objects in the presence of an auditory stimulus such as "more" or "larger", and the smaller of two objects in the presence of "less" or "smaller". Soon, these cues can come to control a generalized pattern of relational responding applicable to any stimuli, no matter what their physical properties; hence the term 'arbitrarily applicable relational responding'. For example, a child with a sufficient history of exposure to the cues "more" and "less", on being told that "A is less than B and B is less than C", will derive (in accordance with a generalised pattern) that "B is more than A", and "C is more than B" (i.e., deriving a relation in the reverse direction from the provided one, sometimes referred to as *mutual entailment*), as well as that "A is less than C", and "C is more than A" (i.e., the combination of provided relations, sometimes referred to as *combinatorial entailment*; e.g., Berens & Hayes, 2007). These emergent outcomes occur despite the lack of any obvious physical comparative relation between A, B and C. The pattern of derivation is based instead on the contextual cues involved. Over recent years, behavior analysts have made substantial progress providing empirical demonstrations of different varieties of AARR (Dymond, May, Munnelly, & Hoon, 2010). Despite this, several key forms of AARR, including spatial relations, remain unexplored (e.g., Stewart, Tarbox, Roche, & O'Hora, 2013).

In accordance with this perspective, spatial relational responding likely develops initially as a purely nonarbitrary based form of relational responding. For example, a child might be taught to respond to stimuli in his or her current context in terms of physical location. For instance, he or she might learn to place objects physically above or below, in front of or behind, or to the left or right of other objects and things. After sufficient training with such physically based nonarbitrary relational responding, the child may derive spatial relations in accordance with spatial cues in the absence of further input from directly available physical properties, that is, show completely abstract or arbitrarily applicable spatial relational responding. As an example, consider someone who responds to the stimuli 'north', 'east', 'south', and 'west' as conventional cues for spatial relational responding. If this person is told that "B is two miles north of A", that "C is one mile east of B", and that "D is two miles south of C", then despite the fact that no absolute physical location has been specified for any of these arbitrary stimuli, the person should derive a variety of untaught spatial relations. For example, s/he should correctly answer questions such as, "If one was at C and wanted to get to B, which direction would one walk?" (in this case, by replying, "West"). S/he should also correctly answer questions requiring knowledge of several spatial relations, such as, "How would one get from A to D?": "Walk east for one mile" can be derived by combining the information provided in the three stated premises.

The aim of this study was to develop and validate the first functional analytic model of spatial reasoning, drawing on the general conception just outlined. Before describing the approach taken in more detail, however, we will first discuss an important distinction between task types employed in examining spatial reasoning that is of relevance to the model to be described. As indicated above, studies investigating spatial TI in adults have often involved exposing participants to a series of verbal premises (e.g., ‘A is to the left of B’, ‘B is to the left of C’) during training, followed by test questions about stimulus pairs that were not explicitly given as one of the premises (e.g., “Where is A relative to C?”) (e.g., Klauer et al., 1997). Meantime, research investigating spatial TI in children has involved the presentation of the physical positions of the premise pairs during training. For example, Andrews and Halford (2002) provided children with a series of five ‘premise towers’ consisting of pairs of different colored blocks. Following training, they were then tested on their ability to infer the relative position of the colors if they were to be placed into a larger block tower. That is, having been provided with a red block (top) yellow block (bottom) pair, and a yellow block (top), blue block (bottom) pair, participants were asked questions such as, “Which color will be higher up in the tower – the red one or the blue one?”, in the absence of any further positional cues. The findings indicated that children as young as four responded correctly to these spatial reasoning tasks. In functional analytic terms, the first (adults) paradigm is an example of AARR alone (i.e., both the trained and derived relations are solely dependent on contextual cues; e.g., ‘left’); or, in more conventional terms, it is purely ‘abstract’. In contrast, the second (children) paradigm involves both non-arbitrary relational responding (i.e., the physical spatial relations used in training) as well as arbitrarily applicable relational responding (i.e., responding to the contextual cues presented in the question with a previously unseen and thus derived relation); in more conventional terms, this is a more ‘concrete’ task. In the current study, we focused primarily on modelling the purely abstract reasoning task, because this is a better representative of the type of reasoning that might be involved in complex reasoning and problem-solving and that predicts intellectual excellence in Science, Technology, Engineering, and Mathematics (STEM) domains (Lubinski, 2010; Wai, Lubinski, & Benbow, 2009). Towards this end we focused on inducing participants to show arbitrarily applicable spatial relational responding in which both the trained and tested relations were based on contextual cues alone. At the same time, however, we also incorporated an element in the protocol that was more representative of the partly non-arbitrary paradigm, specifically, a training phase (Phase 4) introduced to provide 'remedial' training to participants failing the purely AARR stage.

In the current study, we adapted an existing automated protocol, the Relational Completion Procedure (RCP; Dymond & Whelan, 2010), to train and test arbitrarily applicable spatial relational responding in adult humans. The RCP has been employed in previous studies to show performances indicative of TI with comparative relations (Munnelly, Dymond & Hinton, 2010), as well as establishing equivalence (Dymond & Whelan, 2010; Walsh, Horgan, May, Dymond, & Whelan, 2014), and opposition relations (Bennett, Hermans, Dymond, Vervoort, & Baeyens, 2015; Dymond, Roche, Forsyth, Whelan, & Rhoden, 2007; Dymond, Ng, & Whelan, 2013). For the present purposes one important difference between the procedures employed here versus those in previous experiments concerns the nature of the relational response requirement. Previously, participants were required to drag and drop one or both of the *relata* in order to complete a statement involving both the *relata* and the *relation*. For example, Dymond and Whelan (2010) presented the following task in relational training: A OPPOSITE ?, and participants had to drag and drop one of the comparison relata (either the B, C, or D stimulus) into the available space. In the present experiments, participants were taught to drag and drop the specific *relation* contextual cue (cf. Lipkens & Hayes, 2009). For example, when presented with A and B (i.e., the relata), participants had to select one of four comparison stimuli which correctly specified the relationship between them (i.e., *left of, right of, above* or *below*).

Phase 1 of Experiment 1 established contextual cues for the relations, "left of", "right of", "above", and "below" by training participants to select one of four nonsense trigrams depending on the relative positions of two onscreen stimuli. In Phase 2, the contextual cues were used to train and test a novel arbitrary spatial relational network. The network consisted of three directly trained relations, or premise pairs, and three mutually entailed and two combinatorial entailed test relations (see Table 1). Where necessary, remedial training was incorporated to facilitate spatial relational responding, if absent following initial training.

Remedial training involved two changes to the training and testing. First, participants were provided with additional nonarbitrary support during the training of the premise pairs. That is, during these trials the respective stimulus pair appeared on the screen in their corresponding spatial positions. Second, feedback was provided for all responses. These two adaptations were consistent both with the documented facilitative effect of providing nonarbitrary support and feedback for relational responding across multiple exemplars (Barnes-Holmes et al., 2001; Walsh et al., 2014; Berens & Hayes, 2007), and the findings within the broader literature on spatial TI suggesting that providing the positions of the premise pairs during training facilitates correct responding in children (Andrews & Halford, 2002; Pears & Bryant, 1990). Upon completion of remedial training with one set of stimuli, participants were trained and tested with a novel set.

Experiment 2 replicated the procedures from Experiment 1 but extended the analysis by incorporating a reversal component into the research design. Specifically, following successful spatial relational responding, each member of the stimulus set was reassigned to a new network location. Contingent upon criterion-level relational responding in the reassigned network, the participants were then given training and testing with the stimuli in the original network locations.

INSERT TABLE 1 HERE

Experiment 1

Method

Participants

Four experimentally naive participants (3 females, 1 male), aged between 20 and 32 years old, were recruited through personal contacts as well as a university-wide email call for volunteers. Participants did not receive compensation and all provided informed consent.

Setting and Materials

Sessions took place in a small room containing only a table, chair, touch-screen monitor, and computer programmed in Visual Basic.NET© that controlled all stimulus presentations and recorded all responses. The experimenter attended all sessions but sat outside the room. Participants were not permitted to take any belongings into the testing room. P1, P2, and P3 all engaged in a single training/testing session that lasted between 60 and 90 minutes. P4 engaged in two training/testing sessions that lasted approximately 60 minutes each. Four nonsense trigrams (PAF, CUG, VEC, and JOM) were employed as contextual cues for 'Left of', 'Right of', 'Above', and 'Below', respectively. During the nonarbitrary training and testing, stimuli (S1-S6) consisted of six pictures of common objects (see Figure 1) obtained from the Microsoft © Clipart directory. During the arbitrary training and testing, stimuli consisted of eight Wingdings characters. The stimuli were randomly partitioned into two sets of four stimuli. Each stimulus within each set was then assigned to a position in the to-be-trained relational network. This resulted in two stimulus sets consisting of the following set and position designations: Set 1: A1, B1, C1, D1; Set 2: A2, B2, C2, D2. For the purposes of clarity, letters denote the stimulus position in the network (A = top left, B = top right, C = bottom right, and D = bottom left) and the numeral denotes the set number (see Figure 2). All stimuli were approximately 40 mm^2. The trials for nonarbitrary and arbitrary relational training and testing are described using the following convention: The first sample stimulus appears in capitals, followed by the second sample stimulus, and then by the correct comparison stimulus (a contextual cue) which is shown in brackets. For example, the notation B1/C1 [*above*] indicates that in the presence of the sample stimuli B1 (first) and C1 (second), the contextual cue *Above* (rather than one of the alternative cues *Right of*, *Left of*, or *Below*) was designated correct.

INSERT FIGURE 1 HERE

Procedure

Experiment 1 consisted of four phases (see Table 2). Phase 1 involved nonarbitrary relational training (and testing) designed to establish four nonsense syllables as contextual cues for (nonarbitrary) spatial relational responding. Phase 2 employed the cues established in Phase 1 to train an arbitrary spatial relational network. Phase 3 tested for spatial relations that might be derived based on the arbitrary spatial relational network trained in Phase 2. Phase 4 consisted of a remedial training phase designed to facilitate relational responding for participants who had failed Phase 3.

INSERT FIGURE 2 HERE

Programmed Consequences. Feedback for correct responding involved the word “Correct” being displayed for 1 s beneath the selected stimuli. Feedback for incorrect responding involved the word “Incorrect” being displayed for 1 s beneath the selected stimuli. In both cases, the stimuli and the text were presented inside a yellow rectangle. In addition, a progress bar was present at the bottom center of the screen. The progress bar appeared concurrently with the appearance of comparison stimuli and remained onscreen until the onset of the intertrial interval (ITI), during which the entire screen was blank. The progress bar was designed to give participants feedback on how far they had progressed through a phase. During phases in which reinforcement was not available for responses (e.g., Phase 1c and Phase 3) the bar filled incrementally following each participant response regardless of whether a correct or incorrect selection was made. This was explained to participants during pre-phase instructions.

INSERT TABLE 2 HERE

Phase 1: Nonarbitrary relational training. Phase 1 established the four trigrams as contextual cues for left of, right of, above and below. Phase 1 included three stages (1a, 1b and 1c). Stages 1a and 1b involved contextual cue training of increasing complexity. In Stages 1a and 1b, participant responses were followed by feedback (either correct or incorrect) while in Stage 1c, which served as a test phase, feedback was not provided.

Prior to the onset of Phase 1a, participants were presented with the following on-screen instructions:

In a moment a number of pictures will appear on the screen. Two pictures will appear at the top of the screen followed by four nonsense words at each corner. Your task is to drag and drop one of the words into the blank space to complete a statement about the two pictures at the top of the screen. You will be asked to confirm or cancel your selection. There will always be a correct and an incorrect answer. Initially you will be given feedback on your selections, later you will not. Your aim is to get as many correct as possible. The bar at the bottom indicates your progress through the phase of the experiment; it does not indicate whether an answer was correct or incorrect. Please press start if you wish to continue.

Once the participant pressed 'start', stimuli were sequentially presented on the screen. Phase 1a began by training participants in S1/S2 [left of] trials. First, a stimulus (S1) appeared in the top half of the screen slightly to the left of center, followed 1s later by a second stimulus (S2) which appeared to the right of the first stimulus. Following a further 1s delay, all remaining stimuli appeared on screen; the four contextual cues (trigrams) appeared simultaneously in each corner of the screen, along with duplicates of S1 and S2 which appeared either side of a horizontal black line. S1 and S2 were presented in the same order (left to right) in which they appeared at the top of the screen (see Figure 3 for the stimulus presentation sequence). The progress bar appeared at the same time as the four contextual cues. Participants were required to drag and drop one of the trigrams onto the horizontal line shown between S1 and S2. The correct trigram was indicated by two features of the trial presentation; first, the relative spatial positions of S1 and S2 in the top half of the screen (i.e., S1 was physically left of S2); second, the position of S1 and S2 relative to the horizontal line in the bottom half of the screen. The stimulus presented first (in this case S1) had to be judged relative to the stimulus presented second. Once participants had selected the trigram, they were then prompted to confirm their selection. The three remaining trigrams were removed from the screen, and two rectangles, measuring approximately100 mm x 40 mm, appeared side-by-side directly underneath the horizontal line but above the progress bar. The rectangle to the left was colored green and contained the text 'Confirm'. The rectangle to the right was colored red and contained the text 'Cancel'. If participants selected 'Confirm', they were given feedback on their selection, while if they selected 'Cancel', the chosen trigram was removed from the line and repositioned in the respective corner from which it had been dragged. Following the confirmation of a selection, the feedback screen appeared for 1 s. This was followed by a 3 s ITI in which all stimuli were cleared from the screen, before the onset of the next trial. In all training and testing sessions, the order of trials, and the position of the comparison stimuli, was quasi-randomized.

During Stage 1a, S1/S2 [left of] trials were presented in a block of eight trials. Participants were required to meet a criterion of 7 out of 8 correct responses (87%) in order to progress to the next stage of Phase 1a. Upon meeting criterion, they were presented with blocks of eight trials involving S1/S2 [above] trials. Here, S1 appeared in the top half of the screen in the center, followed 1s later by S2, which appeared directly underneath S1. All other aspects of the trial presentation were the same as before. Participants were presented with S1/S2 [above] trials in isolation until they had met a criterion of 7 out of 8 correct responses in one trial block. Upon mastery, participants were given training in S1/S2 [left of] and S1/S2 [above] trials in a mixed block of eight trials in which each trial type was presented four times. Following criterion level responding (7 out of 8 correct responses) in the mixed training, participants began S1/S2 [below] training. Here, S1 appeared in the center of the top half of the screen, followed 1s later by S2, which appeared directly above S1. All other aspects of the trial presentation were administered as before. Training continued until participants demonstrated mastery of 'below' relations (7 out of 8 correct in one trial block) at which point they progressed to mixed training of 'left of', 'above', and 'below' contextual cue training involving a 12-trial blocks of S1/S2 [left of], S1/S2 [above], and S1/S2 [below] trials, in which each trial type was presented four times. Upon mastery of the mixed training (11 out of 12 correct in a single block), participants progressed to training with S1/S2 [right of] in blocks of 8 trials. During S1/S2 [right of] trials, S1 appeared in the top half of the screen slightly to the right, followed 1s later by S2, which appeared to the left of S1. Participants completed Stage 1a when they met the criterion of 87% correct (7 out of 8) in the final training block, which involved a mix of S1/S2 [left of], S1/S2 [above], S1/S2 [below], and S1/S2 [right of] trials, presented twice each.

Upon meeting criterion for Stage 1a, participants progressed to Stage 1b. Stage 1b consisted of a block of mixed trial types identical to the final trial block of Stage 1a with two exceptions. First, S1 and S2 were replaced by two new stimuli: S3 and S4. Second, Stage 1b involved the introduction of trials in which the order of the stimuli in the relational statement was systematically varied. For example, participants were presented with S3/S4 [left of] trials in an analogous fashion to Stage 1a, but were also presented with S4/S3 [right of] trials. Participants met criterion when they had responded at 87% accuracy (7 out of 8 correct) to one block of eight trials containing one presentation of all possible stimulus combinations with S3 and S4 (e.g., S3/S4 [left of], S3/S4 [above], S3/S4 [below], S3/S4 [right of], S4/S3 [left of], S4/S3 [above], S4/S3 [below], and S4/S3 [right of]).

INSERT FIGURE 3 HERE

In the final stage of Phase 1 (1c), participants responded to a block containing the same trial types as encountered during Stage 1b except that S3 and S4 were replaced by two novel stimuli, S5 and S6 (S5/S6 [left of], S5/S6 [above], S5/S6 [below], S5/S6 [right of], S6/S5 [left of], S6/S5 [above], S6/S5 [below], and S6/S5 [right of]). No feedback was provided during Stage 1c. Participants were required to achieve a score of 88% correct (7 out of 8 correct) in order to progress to Phase 2. They were provided with eight opportunities to meet the mastery criterion for each stage before being prompted to contact the experimenter.

Phase 2: Arbitrary relational training. Participants were presented with Phase 2 immediately after reaching criterion with Phase 1. The aim of this phase was to train an arbitrary spatial relational network using the trigrams established as contextual cues in Phase 1. Three relations were trained; A1/B1 [left of], B1/C1 [above], and C1/D1 [right of]. These three relations were selected because this combination represented the minimum number of trained relations required to test for both mutually entailed and combinatorial entailed relations across all four stimuli. Prior to commencing Phase 2 participants were provided with the following instructions.

In this part of the experiment you will be presented with a statement to complete. However, you will have to work out the right answer without the appearance of any additional picture stimuli. Sometimes you will be told whether your selections are correct and other times you will not. There is always a correct and incorrect answer. Try to get all the tasks correct. Press 'start' to continue.

During Phase 2 trials, a screen simultaneously presented the four contextual cues in each corner and two stimuli were presented at the bottom of the screen separated by a horizontal line in an identical fashion to Phase 1. In contrast to Phase 1, however, no additional stimuli appeared on the screen. Participants were required to drag and drop one of the contextual cues onto the horizontal line and to then confirm their selection (see Figure 4 for the stimulus presentation sequence). All responses were followed by feedback. Trial blocks consisted of 12 trials with each relation presented four times. To meet criterion for this phase of training, participants were required to respond correctly to 12 out of 12 trials (100%) within a single block. Failure to achieve criterion resulted in the participant beginning the training block again. Participants were provided with six opportunities to pass Phase 2.

INSERT FIGURE 4 HERE

Phase 3: Arbitrary relations posttest. Upon reaching the Phase 2 criterion, participants were immediately presented with a posttest. The purpose of this phase was to determine whether the arbitrary training provided in the previous phase would result in spatial relational responding. Testing occurred in a block of 32 trials. Twelve of these tested the relations trained in Phase 2 while the remaining 20 tested the following derived relations: B1/A1 [right of], C1/B1 [below], D1/C1 [left of], A1/D1 [above], and D1/A1 [below]. In order to ensure that the trained relations were intact, a criterion of 92% accuracy (11 out of 12) was imposed. If this criterion was not met, the posttest was readministered and participants could be recycled in this manner up to a maximum of six times. If participants passed the criterion for trained relations, yet scored < 90% correct for derived relations (i.e., less than 18 out of 20 correct) then they progressed to Phase 4 (remedial training). If they met criteria for both trained and derived relations then they were deemed to have demonstrated spatial AARR and exited the experiment.

Phase 4: Remedial training. Prior to this phase, the following instructions were presented to participants:

> *In this part of the experiment you will be presented with a statement to complete in the presence of stimuli on the screen but at other times you will have to work out the right answer without the appearance of any additional stimuli. You will be told each time whether you got the answer correct or incorrect. There is always a correct and incorrect answer. Try to get all the tasks correct.*

During Phase 4, trial blocks were presented in an identical fashion to Phase 3 with the following two exceptions. First, participants received feedback for responses to all trials. Second, during trials for A1/B1 [left of], B1/C1 [above], and C1/D1 [right of] (i.e., the relations *trained* in Phase 2) participants were provided with additional on-screen stimuli. Specifically, prior to the presentation of the two sample stimuli in the relational statement, a 2 x 2 grid appeared at the top of the screen in which were placed two stimuli. The stimuli were presented in one of three possible arrangements, which varied as a function of the particular relation targeted (see Figure 4).

During A1/B1 [left of] remedial training trials, A1 appeared in the top left square and B1 appeared in the top right square. During B1/C1 [above] trials, B1 appeared in the top right square and C1 appeared in the bottom right square. Finally, during C1/D1 [right of] trials, C1 appeared in the bottom right square and D1 appeared in the bottom left square. Participants were required to meet the criterion of 92% correct (11 out of 12) for these three trial types. In addition, participants were required to achieve 90% correct (18 out of 20) for the derived relations trials during a single 32-trial block. If participants met the criterion within six exposures to the 32-trial block they were returned to Phase 2 with a new stimulus set (Set 2). Participants who failed to meet criterion within six blocks of remedial training were excused from further participation.

Results and Discussion

Phase 1: Nonarbitrary relational training and testing. Table 3 shows the results of nonarbitrary training and testing. All four participants completed training Phases 1a and 1b within the required number of training blocks and therefore progressed to the test (Phase 1c). P2, P3, and P4 passed this test on their first exposure while P1 required one further exposure before passing. All participants progressed to the arbitrary training phase of the experiment (Phase 2).

INSERT TABLE 3 HERE

Phases 2 to 4: Arbitrary relational training and testing. Table 3 and Figure 5 show the results of the arbitrary training and testing phases. All participants achieved the required mastery criterion during the arbitrary training phase with the initial stimulus set (Phase 2). One participant, P2, passed the subsequent test for derived relations (Phase 3) on the first test exposure (100% accuracy). The remaining three participants failed to meet criterion for derived relations during the initial posttest. P1 responded at 100% accuracy for the trained relations, and 0% accuracy for the derived relations. P3 responded accurately to 11 out of the 12 baseline trials (92%), but at 20% accuracy on the derived relations probes. P4 responded correctly on all baseline relations trials but on only 15% of the derived relations trials. Thus, P1, P3, and P4 progressed to Phase 4 (remedial training). During remedial training (Phase 4), P1 and P4 met criterion following two and four blocks respectively. P3 failed to achieve the mastery criterion following six blocks of training (192 trials) and was therefore excused from further participation. P1 and P4 progressed to arbitrary training and testing with a second set of stimuli. They both met the mastery criterion for the arbitrary relational training (Phase 2). During the subsequent tests for derived relations (Phase 3), both responded at criterion levels for both baseline relations (100% correct) and derived relations (100% correct) on their first test exposures (see Figure 5).

INSERT FIGURE 5 HERE

In summary, three out of four participants passed tests of derived arbitrarily applicable spatial relational responding, one (P2) on their first exposure and two (P1 and P4) following remedial training with a single set of stimuli. The facilitative effect of the remedial training is consistent with the existing literature from the work on multiple exemplar training instruction (Barnes-Holmes et al., 2001; Walsh et al., 2014; Berens & Hayes, 2007) as well as the trend within the broader research literature for including nonarbitrary relational support when testing spatial TI in children or other populations who might be presumed to have less advanced relational repertoires.

These findings represent the first empirical demonstration of arbitrarily applicable spatial relational responding.

Experiment 2

Experiment 1 provided a demonstration of spatial relational responding in three out of four participants, two of whom required the remedial training procedure. The aim of Experiment 2 was to extend on this initial demonstration experiment in two main ways. First, a variant reversal design (e.g., Smyth, Barnes-Holmes, & Forsyth, 2006) was implemented in which the stimuli involved in the initial arbitrary relations training and testing were then retrained and retested in new positions in the relational network before being subsequently retrained and retested once again in the original configuration. This variation on the previous test was designed to demonstrate a more robust level of control over arbitrarily applicable derived spatial relational responding than that shown heretofore.

The other change made in Experiment 2 was to intersperse nonarbitrary training trials with unrelated stimuli during arbitrary training. This was done to increase the likelihood that established contextual cue functions would be maintained throughout testing, which was particularly important given the occasionally long breaks between experimental sessions.

Also, given the large number of trials it took P3 to master Phase 1b during nonarbitrary training in Experiment 1, it was hypothesised that breaking this phase down further into graduated increments would increase the speed at which participants might transition through the nonarbitrary phases from training to testing.

Method

Participants

Thirteen experimentally naive participants (9 females, 4 males), aged between 18 and 59 years old, were recruited via personal contacts and following a university wide email call for volunteers. Participants did not receive compensation and all provided informed consent.

Setting and Materials

All sessions took place in a quiet room containing a table and chairs and a 17" laptop computer programmed in Visual Basic.NET© that controlled all stimulus presentations and recorded all responses. Participants 5, 6 and 7 were trained with arbitrary relations involving the same stimulus sets as used in Experiment 1 (i.e., Wingdings) plus one extra stimulus set (Set 3; see Figure 1). Participants 8 to 17 were trained with abstract colored objects called *Fribbles* (Sets 4 – 8; see Figure 1; Images courtesy of Michael J. Tarr, Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University; http://www.tarrlab.org/).

Procedure

Participants took between 1 and 8 teaching/testing sessions to complete the experiment. Each session lasted between 20 and 120 minutes. During sessions lasting longer than 20 minutes, participants were provided with a 5-minute break between each phase of the experiment (every 10-15 minutes) during which they remained in the testing area. There were a number of changes from Experiment 1. There were seven rather than four phases (see Table 3). Phase 1 consisted of nonarbitrary relational training (and testing). Phase 2 involved a pretest of arbitrary relations. Phase 3 trained an arbitrary spatial relational network. Phase 4 tested for the emergence of spatial relational responding. Phase 5 consisted of the remedial training phase designed to facilitate spatial relational responding in participants who had failed Phase 4. Phase 6 consisted of a reassignment training and testing phase in which the same set of stimuli with which participants had successfully derived relations were reordered into novel positions as part of a new relational network. Finally, Phase 7 functioned as a reversal phase in which participants were retrained and tested with the same relational network configuration as in Phase 3.

Programmed Consequences. Reinforcement was not available for any response during Phase 1d, or for targeted derived relations during Phase 2, Phase 4, Phase 6 and Phase 7. In all such cases, however, participants were informed that the on-screen progress bar filled incrementally on each trial regardless of whether a correct or incorrect selection was made.

Phase 1: Nonarbitrary relational training. Two changes were made to the Phase 1 training during Experiment 1. Firstly, Stage 1b involved the presentation of a block of eight trials identical to the final block of Phase 1a but with two new stimuli (S3 and S4). This differed from Experiment 1 in which both the stimuli set and the trial types were changed simultaneously. Thus, Stage 1b involved trials consisting of S3/S4 [left of], S3/S4 [above], S3/S4 [below], and S3/S4 [right of] trials, presented twice each. Secondly, Stage 1c involved the gradual introduction of trials of increasing complexity in which the position of the sample stimuli at the top of the screen remained the same, but the order in which they appeared at the bottom (the relational statement) varied. For example, in the first block of Stage 1c, participants were presented with four trials of S4/S3 [above] and four trials of S3/S4 [below]. Trial types were gradually introduced across trial blocks (see Table 4) until, ultimately, participants responded at 87% accuracy (7 out of 8 correct) to a block of trials containing all possible stimulus combinations with S3 and S4 (e.g., S3/S4 [left of], S3/S4 [above], S3/S4 [below], S3/S4 [right of], S4/S3 [left of], S4/S3 [above], S4/S3 [below], and S4/S3 [right of]). This differed from Experiment 1 in which all trial types (left of, right of, above, and below) were included in the same training block from the outset. In the final stage (1d) of the phase, participants responded to an identical trial block to that encountered in the final block of Stage 1c except that S3 and S4 were replaced with S5 and S6 respectively and no feedback was provided.

INSERT TABLE 4 HERE

Phase 2: Arbitrary relations pretest. Phase 2 in this experiment involved a pretest of arbitrary relations prior to the implementation of arbitrary relational training which was scheduled for the next phase. Prior to beginning this phase of the experiment, participants were presented with the following instructions:

> *In this part of the experiment you will be presented with a statement to complete, however, you will have to work out the right answer without the appearance of any additional picture stimuli. Sometimes you will be told whether your selections are correct and other times you will not. There is always a correct and incorrect answer. Try to get all the tasks correct. Press 'start' to continue.*

During Phase 2, trials were presented in an identical fashion to Phase 3 from Experiment 1. The following five relation types were presented in the absence of any feedback: B1/A1 [right of], C1/B1 [below], D1/C1 [left of], A1/D1 [above], and D1/A1 [below]. Phase 2 was presented in one block of 20 trials during which each relation was presented four times. No feedback was provided for any trial during Phase 2. Following completion of Phase 2, participants immediately progressed to Phase 3 in the absence of any further instructions.

Phase 3: Arbitrary relational training. This phase was identical to Phase 2 from Experiment 1 but incorporated one additional component; non-arbitrary training trials from Phase 1 were interspersed in the training blocks (see Table 4). This adaptation helped to ensure that the trigrams established as contextual cues for particular patterns of spatial relational responding retained their respective functions. All responses were followed by feedback. Trial blocks consisted of 20 trials in total; each arbitrary relation was presented four times (12 trials) alongside four non-arbitrary relations, which were presented twice each (8 trials). To meet criterion for this phase of training participants were required to respond correctly on all arbitrary trials, and on 7 out of 8 nonarbitrary trials (87%) within a single block of trials. Failure to achieve either criterion resulted in the participant being re-exposed to the training block. This cycle was repeated up to a maximum of six times.

Phase 4: Arbitrary relations posttest. This phase was identical to Phase 3 from Experiment 1.

Phase 5: Remedial training. This phase was identical to Phase 4 from Experiment 1 with two exceptions. First, upon meeting the criterion for a remedial training trial block participants were returned to Phase 2 with a second set of stimuli. Second, a maximum of four additional sets of stimuli were available for training during remedial training. If participants passed Phase 4 with Set 2 they immediately progressed to Phase 6 with that set. If they failed Phase 4 with the second set, they were provided with additional remedial training, returning to Phase 2 with a third set of stimuli. This cycle of train, test, and remedial training continued for up to four sets of stimuli. Participants who failed to meet criterion within six blocks of remedial training with any of the four stimulus sets were excused from further participation in the experiment.

Phase 6: Reassignment training and testing. During this phase, participants were trained in a new relational network with the same stimulus set with which they had passed

Phase 4 but with the stimuli re-arranged in order to ensure that all stimulus relations differed from those originally trained and tested. In practice, this involved the following network adaptations: stimulus A replaced stimulus C, stimulus B remained in the same position, stimulus C replaced stimulus D, and stimulus D replaced stimulus A (See Table 4). Participants underwent training and testing with this reassigned network analogous to that received during their original exposure to Phases 3 and 4 respectively. Phase 6a corresponded to the training (Phase 3) and Phase 6b to the posttest (Phase 4). If they met criterion during the posttest (Phase 6b), they progressed to Phase 7. If, however, they failed to meet criterion following six exposures to either the training or the test blocks, they were excused from further participation.

Phase 7: Reversal. Phase 7 involved re-establishing the relational network as originally trained during Phase 3, and tested during Phase 4. Thus, Phase 7 employed training and testing identical to that presented during these phases. Phase 7a corresponded to the training (Phase 3), and Phase 7b to the posttest (Phase 4). Upon meeting criterion in the posttest, participants exited the experiment; if they failed to meet criterion following six exposures to the training or testing blocks they were excused from further participation.

Results and Discussion

Phase 1: Nonarbitrary relational training and testing. Eleven out of the thirteen participants in Experiment 2 passed the training and testing criteria for Phase 1; the exceptions were P11 and P15 who failed to meet criterion for Stage 1c following 80 (P11) and 128 (P15) trials. Following P11 and P15's failure to achieve criterion, they were provided with 'adapted' non-arbitrary training. For P11, this involved repeating Phases 1a-1c; however, during trials, an arrow specifying the direction of the to-be-trained relation (i.e., left, right, up, or down) accompanied the sample stimuli. Following adapted training, P11 progressed to the non-arbitrary test phase (Phase 1d) and met criterion on the first test exposure. P15 was provided with the same adapted non-arbitrary training but in contrast to P11 this occurred at Phase 1c. Following an additional 96 trials in which P15 failed to meet the passing criterion, this participant opted to exit the experiment. Of the eleven participants that passed the standard non-arbitrary training and testing sequence, P5, P9, and P10 underwent two exposures to the test. In each case, participants met criterion during the initial test exposure; the second test administration occurred due to a pause in participation in the study (i.e., a session break of more than 24 hours). The readministration of the test was conducted (Phase 1d) in order to ensure that the functions for the contextual cues had been maintained prior to the implementation of Phase 2. Table 5 shows the number of trials to criterion during each phase of training and testing.

INSERT TABLE 5 HERE

Phase 2-5: Arbitrary relational training and testing. Eight out of the twelve participants who progressed to Phase 2 (the pretest phase), subsequently met the criterion for the arbitrary relational training in Phase 3 (viz., P5, P8, P9, P10, P11, P14, P16, and P17). Seven out of the eight participants who progressed to the arbitrary relational testing (Phase 4) passed the tests for derived relations (all participants except P14). Figure 6 shows the results of all pretests and posttests. With the exception of P5 and P16, all participants passed tests for derived relations with Set 2, following the implementation of remedial training with Set 1. P5 passed the test for derived relations with Set 3 following remedial training with two stimulus sets. P16 met criterion for derived relations during the pretest for Set 2, so underwent a further pretest (and subsequent training) with Set 3. P5, P8, P9, and P17 passed the tests on the first posttest exposure (Set 2) following remedial training. P16 also passed the tests on the first posttest exposure (Set 3) following remedial training. P10 met criterion for derived relations (Set 2) on the third test exposure having failed to meet the trained relations criteria on the first (10 out of 12 correct) and second (10 out of 12 correct) test blocks. P11 initially failed to meet criterion during remedial training (Phase 5) following six cycles of training (192 trials). This resulted in the implementation of an adapted remedial training phase. During this adapted training, all four stimuli (rather than two) from the relational network appeared in their respective positions in the on-screen grid during the directly trained trials. All other aspects of the remedial training remained the same. P9 did not receive a pretest for Set 1 due to experimenter error.

Four participants P6, P7, P12, and P13, failed to meet the arbitrary relational training criterion (Phase 3) with Set 1. P6, P7, and P12 underwent six cycles of training (120 trials) before exiting the experiment. Following six cycles of training, P13 elected to continue the training; however, he failed to meet the criterion following an additional 120 trials and was excused from further participation. One participant, P14, met the required training criterion, but failed to pass tests for derived relations despite undergoing remedial training with four sets of stimuli. P14 required three test exposures with Set 2 following a failure to maintain the baseline relations. During the third test block P14 responded at 60% accuracy for derived relations. During testing for Sets 3, 4 and 5, he scored 50%, 50%, and 55% correct, each during the first test exposure for derived relations. Having failed to demonstrate the required pattern of responding following remedial training with four sets of stimuli, P14 exited the experiment.

INSERT FIGURE 6 HERE

Phases 6 - 7: Reassignment and reversal. All seven participants that passed tests for derived relations (Phase 5) progressed to training and testing for a reassigned relational network. Six participants met criterion during the test for derived relations (Phase 6b). P5, P8, P11, P16 and P17 passed on the first posttest exposure, while P10 passed following two posttest exposures. P9 failed to meet criterion for the derived relations during the posttest following reassignment training. P9 obtaining a score of 70% correct for the derived relations, at which point P9 exited the experiment due to time constraints. All of the remaining six participants except P16 also subsequently passed derived testing (Phase 7b) following retraining of the original relational network (Phase 7a). P16 scored just below the required criterion at 92% correct (11 out of 12) for the trained relations and 85% correct (17 out of 20) for the derived relations in the reversal posttest.

In summary, in the current experiment, seven out of thirteen participants passed tests of spatial relational responding, six of whom showed not just the same basic pattern as Experiment 1, but in accordance with a variant reversal design. This variation on Experiment 1 arguably shows a more robust level of control over derived spatial relational responding (i.e., at the individual participant level) than was previously shown.

General Discussion

The present study demonstrated the first functional analysis of spatial reasoning as arbitrarily applicable spatial relational responding. In Experiment 1, three out of four adult participants passed tests of arbitrarily applicable spatial relational responding, one on their first exposure and two after remedial training. In Experiment 2, seven out of the eight participants who met the criteria for the directly trained arbitrary relations also passed the spatial relations tests following remedial training, six of whom showing not just the same basic pattern as in the first experiment, but illustrating control via a reversal design in which the stimuli involved in relations trained and tested in the initial arbitrary training were then retrained and retested in new positions in the relational network, before being subsequently retrained and retested once again in the original configuration. The latter demonstration arguably constitutes a more convincing example of intra-individual control over spatial relational responding than that seen in Experiment 1.

Developmental and cognitive psychologists have amassed a substantial research literature investigating patterns of spatial reasoning (e.g., Gazes, Hampton, & Lorenco, 2015; Vasconcelos, 2008). In contrast, this potentially important area of investigation has thus far been afforded relatively little attention by behavior analysts. The experiments reported here suggest a role for AARR - a functionally defined generalized operant response class – in accounting for this complex repertoire. Given appropriate contextual control individuals trained in particular spatial relations will respond in accordance with a variety of entailed spatial relations with previously unencountered stimuli. Indeed, this is what has been demonstrated in this study and thus the data reported here represent a behaviorally consistent process for spatial reasoning.

Our findings also further advance research on multiple derived stimulus relations by providing the first empirical evidence of a new, complex pattern of relational responding. In the present study, participants showed a type of relational derivation unique in empirical work on derived relations (at least thus far), in which after being trained on three different patterns of (spatial) relations, they responded in accordance with a relation that was based on the combination of the first three but was different from any of those trained. This is a relatively complex pattern of derivation made possible because spatial relations can be specified along more than one dimension. In the current experiments, two dimensions (i.e., X and Y) were employed, though additional dimensions could at least in theory be added (e.g., *n*-dimensional space in mathematics) to introduce even greater derivational complexity. Interestingly enough, despite the use of additional training support, including training of the nonarbitrary relations, it was difficult to induce spatial relational responding even along two dimensions; perhaps derivation along further dimensions might be beyond many individuals. Perhaps the ability to derive relations of this type, however, might be a particularly good test of a spatial reasoning repertoire. In any event, the exploration of spatial arbitrarily applicable relational responding might contribute to more in-depth, functional analytic based understanding of such repertoires (De Houwer, 2011; Hughes, Barnes-Holmes & Vahey, 2012).

Across the two experiments, of the twelve participants that progressed to the remedial training phase (Phase 4), ten passed the subsequent tests for derived relations. What is unclear from the current analysis is which aspects of the remedial training were critical for the emergence. In both experiments, the training involved the provision of reinforcement for the required pattern of relating, as well as the addition of nonarbitrary components (i.e., providing stimulus positions during baseline trials). The addition of the latter is supported at both a theoretical and an empirical level. As described in the introduction, from the current perspective, nonarbitrary relational responding precedes and supports arbitrarily applicable relational responding; hence incorporating additional nonarbitrary relational training into a protocol might make the emergence of appropriate arbitrarily applicable relational responding more likely. Furthermore, findings from previous research are consistent with this (Berens & Hayes, 2007). At the same time, it is certainly possible that the provision of reinforcement for AARR alone, without nonarbitrary cues, might have been sufficient to produce generalized patterns of relating. Two features of the present data render this latter suggestion unlikely however. First, despite acquiring the patterns of nonarbitrary relational responding during Phase 1 of the experiments, many participants had difficulty mastering the arbitrary training during Phase 2 in which nonarbitrary support was unavailable. Second, P11's data indicate that increasing nonarbitrary support during the adapted training procedure served a facilitative effect; P11 passed the remedial training following the introduction of nonarbitrary stimuli as part of both trained *and* derived trials. Future research should examine this issue more closely, however.

A related conceptual consideration is whether the remedial training could be characterized as an example of multiple exemplar instruction (MEI). MEI is proposed to be a key mechanism in accounting for the emergence of relational operants, and other generalized operant response classes, such as generalized imitation (e.g., Baer & Deguchi, 1985; Catania, 2012). MEI usually involves the application of a reinforcement contingency across multiple examples of a specified response pattern. Typically, the contingency is applied until such time as the same pattern of responding emerges (or generalizes) to novel stimuli in the absence of reinforcement. Certainly, the training protocol employed in the current experiments involved reinforcement of multiple exemplars; however, as discussed above, other aspects of the trial presentations during remedial training also differed. During baseline trials (i.e., trained relations) of the remedial training, participants were provided with the positions of the stimuli (i.e., the nonarbitrary relational support described above) as well as an additional 2 x 2 grid. The inclusion of these additional features to the baseline trials during remedial training might call into question the suitability of the term MEI given its broader use in the literature.

Experiment 2 provides a successful demonstration of stimulus ‘reassignment’ whereby, in the case of seven out of eight participants, stimuli involved in relations trained and tested as part of an initial arbitrary spatial relational network were then retrained and retested in new positions in that network, before being subsequently retrained and retested once again in the original configuration. The successful demonstration of reassignment in this case differs from some patterns seen in some previous studies involving attempted re-assignment of stimuli in the context of derived equivalence relations, in which, following stimulus re-assignment in baseline training, some patterns of derived relations (i.e., symmetry) followed the re-assignment training while others (i.e., transitivity and equivalence) followed the original baseline training (see, e.g., Pilgrim & Galizio, 1990). There were a number of differences between the basic paradigm involved in the current study and that involved in these previous studies that may have given rise to this difference in outcome. For example, the presentation of a restricted number of stimulus pairs during the tests for reassigned relations in the present experiment, naturally limited the extent to which participants could respond in accordance with the original relational network; during reassignment testing, participants were rarely presented with a particular stimulus pair encountered during the original testing. Other potential factors include the patterns of derived relations involved (spatial versus equivalence), the use of explicit contextual cues in the current study, the use of a non-arbitrary relational training phase in the current study, and the use of remedial training in the current study. Any one of these factors or a combination of more than one of them may have been responsible for the difference in patterns. Future work could attempt to isolate the critical variable or variables through their systematic manipulation.

There are some aspects of the present approach to spatial relational responding that warrant further investigation. First, the experiments were conducted with adults who presumably already had a history of learning involving relatively complex derived relational responding. As with other analyses that have investigated complex relational responding in adult populations (e.g., Griffee & Dougher, 2002; Slattery & Stewart, 2014), the purpose of the present study was to provide a functional analytic model of the repertoires involved in spatial relational responding with the overarching aim of elucidating the variables that might be important in its development (e.g., nonarbitrary relational responding). However, as has been noted in analogous preliminary demonstrations (e.g., Slattery & Stewart, 2014), additional research will be needed both to flesh out the empirical analysis as well as to allow for advances into applied domains.

There are many areas in which the present work could be usefully extended. In Experiment 2, for example, while a sizeable number of participants showed the final performance, a number failed at an earlier stage. A high proportion of participants (33%)

exited the experiment because they failed to meet criterion during the arbitrary training phase (Phase 3). The arbitrarily imposed limit of six blocks of training (120 trials) may have been an important variable in this respect. To increase the percentage of participants that acquire the requisite repertoire, additional studies might incorporate a simple-to-complex training via the gradual introduction of different trial types or include nonarbitrary support from the outset of training. These suggestions notwithstanding, it should also be considered that the behavior required was relatively complex and given constraints on time and remuneration for adult participants in the context of experimental work of this kind, a certain attrition rate is difficult to avoid (Drake & Wilson, 2008).

Future studies might also explore the emergence of spatial AARR across childhood. The research in the current study was primarily concerned with whether adults can potentially show this performance. As suggested above, an important wider consideration is that the participants would have begun the experiment with a rich history of relational responding. Hence, the contextually controlled relational responding seen in this study would have been based to some extent on this prior history of relating. More precise scientific knowledge of the processes responsible for producing a repertoire of spatial AARR will require the study of individuals without such learning histories. Such work could, for example, explore questions concerning the trajectory of emergence; to what extent it is supported by other forms of relational responding; and how it might be trained in those, either typically or non-typically developing, with particular deficits. There are some indications that relational flexibility may be predictive of scores on intelligence quotient (IQ) measures (O'Hora et al., 2005), and that directly targeting and strengthening generalised relational operants may have a range of educational benefits (Cassidy, Roche, Colbert, Stewart, & Grey; 2016; Cassidy, Roche & Hayes, 2011; Stewart et al., 2013). As such, it seems possible that training spatial relational responding might help improve intellectual functioning in general. In addition, however, training this variety of relational responding might perhaps also especially boost particular repertoires such as spatial reasoning or navigational ability.

Future work might examine the relative influence of the Relational Completion Procedure-variant employed here to train and test spatial relations. Dymond and Whelan (2010) suggest that the Left-to-right (LTR) response requirement of the RCP may serve to mimic the relational processes involved in reading and completing sentences in LTR languages. All the participants involved in the present study were fluent English speakers with extensive histories of reading and completing sentences. The extent to which these pre-experimental experiences facilitated the emergence of relational responding in the present experiments is unknown. A further consideration concerns whether the LTR response requirement interfered with the development of stimulus control during nonarbitrary training. For example, during S1S2 [right of] nonarbitrary training trials, responding was deemed correct if the trigram for 'right of' was selected, despite the fact that, in the lower portion of the screen, S1 appeared to the left of S2. These trials could be termed spatially incongruent relative to S1S2 [left of] trials in which S1 was presented to the left of S2 in both the upper and lower portions of the screen. While this feature of the training may have exerted competing stimulus control with regard to the training of the contextual cues, we predicted that in the context of training all four nonarbitrary spatial relations this interference would be diluted, and thus the training would be successful. The data appear to bear this out; fifteen out of the seventeen participants across the experiments met criterion during the nonarbitrary test phase. Further research may look to investigate the comparative efficacy of other training and testing paradigms, particularly those frequently used in the derived stimulus relations literature (e.g., Sidman, 1994; Stewart, Barnes-Holmes, & Roche, 2004).

Finally, further work might also look at how spatial relational responding varies as a function of the complexity of the relation tested. One common finding reported in the transitive inference literature, and referred to as the Symbolic Distance Effect (SDE; Moyer & Bayer, 1976), is that there is an increase in response accuracy and response speed, as a function of increasing the number of intervening 'nodes' between test stimuli. For example, following training in a 6-term series (AB, BC, CD, DE, EF), the SDE would predict that performance would be superior (faster and more accurate) in tests for derived relations involving BE (two intervening nodes) relative to a test involving BD (one intervening node; e.g., Bryant & Trabasso, 1971; Libben & Titone, 2008; Vasconcelos, 2008). Exploring whether this well reported effect extends to derived spatial relational responding would be salutary (see also, Munnelly, Dymond, & Hinton, 2010).

Investigating the SDE using the procedures described in the present study might also enable behaviour analysts to explicitly test predictions made by cognitive accounts of such phenomena. For example, the SDE has been frequently cited as evidence for the causal influence of mental representations in transitive inference problem solving. The representation account postulates that an individual solving the task constructs a mental image or line involving each of the training stimuli which are represented in the "spatial" order in which they were trained (Johnson-Laird, 1972). This image is then employed to infer the relationship between any pair of stimuli not directly encountered during training. The occurrence of the characteristic effect during testing is purported to provide evidence for the hypothesis; the closer the two test items in the mental image, the harder it will be to determine their relative positions (Acuna, Sanes & Donoghue, 2002). Further work employing the experimental preparations described here might usefully test this account. Consider, for example, the training of two different relational networks. First, a 6-term series which in terms of spatial position, appears to double back on itself (e.g., A left of B, B left of C, C above D, D right of E, E right of F) and second, a 6-term series that produces end points that are, relatively speaking, located at a greater distance from one another (e.g., A left of B, B above C, C left of D, E right of D, F right of E). Such training would establish two networks in which the number of nodes and spatial cues is consistent, but the proximity of the end points (in representational terms) varies. The speed and accuracy of responding to the end pairs (AE) in both networks might then be compared. According to the cognitive proposals described above, a relative advantage would be predicted for the network in which the end points are located at a greater distance from each other. Evidence to the contrary might call into question the utility of such image-based accounts. As suggested, using the procedures described here might help shed new light on these phenomena. In addition, variables theorised by relational frame theorists to be important in predicting fluency of derivation of relations more generally (see, e.g., Hughes et al., 2012) might supplement such novel analyses; these might include, for example, level of experience in deriving in accordance with particular patterns of relational responding (e.g., spatial) or complexity of the relational network involved in a particular task (measured in terms of the number and variety of relations involved; for example, the fact that spatial relational responding includes several sub-types of framing in accordance with different spatial dimensions may make spatial tasks more difficult in some circumstances).

Although we approach the topic of spatial reasoning from the perspective of Relational Frame Theory (RFT), the present findings might also be interpreted from alternative theoretical positions. There is ongoing debate within behavior analysis as to the value of relational responding as an explanatory concept (Moore, 2009; Palmer, 2004). In particular, it has been argued that the patterns of behavior suggestive of relational responding in studies of derived stimulus relations might depend upon other unmeasured, possibly covert, verbal behavior. Thus, relational stimulus control in the present experiments might have been established in the following way: upon seeing the presentation of two stimuli on the screen during nonarbitrary training, participants covertly produced supplementary stimuli such as, "the monkey is to the left of the tennis ball" and "left of is PAF", all which determined the final response (i.e., the stimulus selection). From a mediational perspective, the terminal performance of the subject cannot be fully understood, and therefore accounted for, without reference to the additional problem solving. The empirical work employing the premises and terminology of joint control (e.g., Lowenkron, 1989) and the naming hypothesis (e.g., Petursdottir, Carp, Peterson, & Lepper, 2015) might be usefully adapted and applied to the challenge of the current topic. In addition, arranging methodological preparations capable of evaluating the contribution of covert verbal behavior might proceed by identifying invariant collateral responses. Investigating supplementary measures such as eye tracking (Dube, Balsamo, Fowler, Dickson, & Lombard, 2006), response latencies (Barnes-Holmes et al., 2005), and neurophysiological measures (e.g., Hinton, Dymond, von Hecker & Evans, 2010; Wang & Dymond, 2013) might help shed further light on such issues. A RFT counterargument is that substantial practical progress can and is being made by manipulating contextual determinants of relational responding without theoretical reliance on mediating processes (see for example, Cassidy et al., 2011; Dymond & Roche, 2013; Hayes & Stewart, 2016). Nevertheless, we enthusiastically encourage behavior analysts from the full range of conceptual and theoretical perspectives to engage with hitherto unexplored research domains such as the current one.

In conclusion, the work reported here provides the first functional-analytic empirical investigation of spatial reasoning. This phenomenon, including in particular spatial transitive inference, has traditionally received more research attention from cognitive-developmental psychology than behavior analysis. However, given that the fact that it appears to be implicated in many complex reasoning and problem solving skills it would seem to be an important repertoire for behavior analytic science to study. Developing procedures for assessing, establishing and strengthening spatial reasoning in both typical and atypical populations is an important goal and the work reported here should facilitate additional progress.

References

Acuna, B. D., Sanes, J. N., & Donoghue, J. P. (2002). Cognitive mechanisms of transitive inference. *Experimental Brain Research, 146,* 1-10. doi:10.1007/s00221-002-1092-y

Andrews, G., & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology, 45,* 153-219. doi.10.1016/S0010-0285(02)000026

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2001). Exemplar training and a derived transformation of function in accordance with symmetry. *The Psychological Record, 51*, 287–308.

Baer, D. M., & Deguchi, H. (1985). Generalized imitation from a radical-behavioral viewpoint. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 179 –217). New York, NY: Academic Press.

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. doi.10.1901/jaba.2007.7-06

Bennett, M., Hermans, D., Dymond, S., Vervoort, E., & Baeyens, F. (2015). From bad to worse: Symbolic equivalence and opposition in fear generalization. *Cognition & Emotion, 29, 1137-1145.* doi:10.1080/02699931.2014.973833

Bryant, P. E. (1974). *Perception and understanding in young children*. London: Methun.

Bryant, P. E., & Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature, 232,* 456-458. doi.10.1038/232456a0

Catania, A. C. (2012). *Learning* (5th ed.). New York: Sloan.

Cassidy, S., Roche, B., Colbert, D., Stewart, I., & Grey, I. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learning and Individual Differences, 47,* 222-235. doi.10.1016/j.lindif.2016.03.001

Cassidy, S., Roche, B. & Hayes, S. C. (2011). A relational frame training intervention to raise intelligence quotients: A pilot study. *The Psychological Record*, *61*, 173-198.

De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science, 6*, 202–209. doi.10.1177/1745691611400238

Drake, C. E., & Wilson, K. G. (2008). Instructional effects on performance in a matching-to-sample study. *Journal of the Experimental Analysis of Behavior*, 89, 333-340. doi.10.1901/jeab.2008-89-333

Dube, W. V., Balsamo, L. M., Fowler, T. R., Dickson, C. A., & Lombard, K. M. (2006). Observing behavior topography in delayed matching to multiple samples. *The Psychological Record*, *56*, 233–244.

Dymond, S., May, R. J., Munnelly, A., & Hoon, A. E. (2010). Evaluating the evidence based for relational frame theory: A citation analysis. *The Behavior Analyst, 33*, 97-117.

Dymond, S., Ng, T., & Whelan, R. (2013). Establishing arbitrarily applicable relations of same and opposite with the relational completion procedure: Selection-based feedback. *The Psychological Record, 63*, 1-20. doi.10.11133/j.tpr.2013.63.1.009

Dymond, S., & Roche, B. (2013). *Advances in relational frame theory: Research & application.* Oakland, CA: New Harbinger

Dymond, S., Roche, B., Forsyth, J. P., Whelan, R., & Rhoden, J. (2007). Transformation of avoidance response functions in accordance with same and opposite relational frames. *Journal of the Experimental Analysis of Behavior*, *88*, 249–262. doi.10.1901/jeab.2007.22-07

Dymond, S., & Whelan, R. (2010). Derived relational responding: A comparison of matching to sample and the relational completion procedure. *Journal of the Experimental Analysis of Behavior, 94,* 37-55. doi.10.1901/jeab.2010.94-37Gazes, R. P., Hampton, R.

R., & Lorenco, S. F. (2015). Transitive inference of social dominance by human infants. *Developmental Science*. Doi.10.1111/desc. 12367

Gazes, R. P., Lazareva, O. F., Bergene, C. N., & Hampton, R. R. (2014). Effects of spatial training on transitive inference performance in human and rhesus monkeys. *Journal of Experimental Psychology: Animal Learning and Cognitive, 40*, 477-489. doi:10.137/xan0000038

Griffee, K., & Dougher, M. J. (2002). Contextual control of stimulus generalization and stimulus equivalence in hierarchical categorization. *Journal of the Experimental Analysis of Behavior, 78,* 433–447. doi.10.1901/jeab.2002.78-433

Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences, 14*, 497–505. doi.10.1016/j.tics.2010.08.005

Hayes, J., & Stewart, I. (2016). Comparing the effects of derived relational training and computer coding on intellectual potential in school-age children. *British Journal of Educational Psychology*. doi.10.1111/bjep.12114

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York: Plenum Press.

Hayes, S. C., & Brownstein, A. J. (1986). Mentalism, behavior-behavior relations, and a behavior-analytic view of the purposes of science. *The Behavior Analyst, 9*, 175-190.

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 language and cognition*. (pp. 21 -49). New York: Plenum Press.

Hinton, E., Dymond, S., Von Hecker, U., & Evans, C. J. (2010). Neural correlates of relational reasoning and the symbolic distance effect: involvement of parietal cortex. *Neuroscience, 168,* 138-148. doi.10.1016/j.neuroscience.2010.03.052

Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition. *Journal of Contextual Behavioral Science, 1,* 17-38. doi.10.1016/j.jcbs.2012.09.003

Johnson-Laird, P. N. (1972). The three-term series problem. *Cognition, 1,* 57-82. doi.10.1016/0010-0277(72)90045-5

Klauer, K. C., Stegmaier, R., & Meiser, T. (1997). Working memory involvement in propositional and spatial reasoning. *Thinking and Reasoning, 3*, 9–47. doi. 10.1080/135467897394419

Knauff, M., Strube, G., Jola, C., Rauh, R., & Schlieder, C. (2004). The psychological validity of qualitative spatial reasoning in one dimension. *Spatial Cognition and Computation*, 4, 167-188. doi.10.1207/s15427633scc0402_3

Krumnack, A., Bucher, L., Nejasmic, J., Nebel, B., & Knauff, M. (2011). A model for relational reasoning as verbal reasoning. *Cognitive Systems Research*, *11*, 377–392. doi.10.1016/j.cogsys.2010.11.001

Libben, M., & Titone, D. (2008). The role of awareness and working memory in human transitive inference. *Behavioural Processes, 77,* 43-54. doi.10.1016/j.beproc.2007.06.006

Lipkens, R., & Hayes, S. C. (2009). Producing and recognizing analogical relations. *Journal of the Experimental Analysis of Behavior, 91,* 105–126. doi.10.1901/jeab.2009.91-105

Lowenkron, B. (1989). Instructional control of generalized relational matching to sample in children. *Journal of the Experimental Analysis of Behavior*, *52,* 293–309. doi.10.1901/jeab.1989.52-293

Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences, 49,* 344-351. doi:10.1016/j.paid.2010.03.022

Moore, J. (2009). Some thoughts on the relation between derived relational responding and verbal behavior. *European Journal of Behavior Analysis*, *10*, 31-47. doi.10.1080/15021149.2009.11434307

Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology, 8,* 228-246. doi.10.1016/0010-0285(76)90025-6

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. doi:10.1016/j.beproc.2010.05.007

Nejasmic, J., Bucher, L., & Knauff, M. (2015). The construction of spatial mental models - A new view on the continuity effect. *Quarterly Journal of Experimental Psychology, 68,* 1794-812. doi.10.1080/17470218.2014.991335

O'Hora, D., Pelaez, M., Barnes-Holmes, D., & Amesty, L. (2005). Derived relational responding and human language: Evidence from the WAIS-III. *The Psychological Record, 55,* 155-174.

Palmer, D. C. (2004). Data in search of a principle: A review of relational frame theory: A post-Skinnerian account of human language and cognition. *Journal of the Experimental Analysis of Behavior*, *81,* 189–204. doi.10.1901/jeab.2004.81-189.

Pears, R., & Bryant, P. (1990). Transitive inferences by young children about spatial position. *British Journal of Psychology, 81,* 497–510. doi.10.1111/j.2044-8295.1990.tb02375.x

Petursdottir, A. I., Carp, C. L., Peterson, S. P., & Lepper, T. L. (2015). Emergence of visual-visual conditional discriminations following intraverbal training. *Journal of the Experimental Analysis of Behavior, 103,* 332-348. doi.10.1002/jeab.136.

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

Smyth, S., Barnes-Holmes, D., & Forsyth, J.P. (2006). A derived transfer of simple discrimination and self-reported arousal functions in spider fearful and non-spider-fearful participants. *Journal of the Experimental Analysis of Behavior*, *85*, 223–246. doi:10.1901/jeab.2006.02-05

Slattery, B., & Stewart, I. (2014). Hierarchical classification as relational framing. *Journal of the Experimental Analysis of Behavior, 101*, 61-75. doi.10.1002/jeab.63

Stewart, I. (2016). The fruits of a functional approach for psychological science. *International Journal of Psychology, 51,* 15-27. doi:10.1002/ijop.12184

Stewart, I., Barnes-Holmes, D., & Roche, B. (2004). A functional-analytic model of analogy using the relational evaluation procedure. *The Psychological Record, 54,* 531-553.

Stewart, I., & McElwee, J. (2009). Relational responding and conditional discrimination procedures: An apparent inconsistency and clarification. *The Behavior Analyst, 32*, 309–318.

Stewart, I., Tarbox, J., Roche, B., & O'Hora, D. (2013). Education intellectual development and relational frame theory. In S. Dymond & B. Roche. (Eds.). *Advances in relational frame theory: Research and application*, (pp. 178-198). New Harbinger: CA, USA.

Vasconcelos, M. (2008). Transitive inference in non-human animals: An empirical and theoretical analysis. *Behavioural Processes, 78,* 313-334. doi.10.1016/j.beproc.2008.02.017

Walsh, S., Horgan, J., May, R., Dymond, S., & Whelan, R. (2014). Facilitating relational framing in children and individuals with developmental delay using the relational completion procedure. *Journal of the Experimental Analysis of Behavior, 101,* 51-60. doi.10.1002/jeab.66

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over fifty years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101,* 817-835. doi.10.1037/a0016127

Wang, T., & Dymond, S. (2013). Event-related potential correlates of emergent inference in human arbitrary relational learning. *Behavioural Brain Research, 236*, 332-343. doi:10.1016/j.bbr.2012.08.033

Figure 1

Experiment 1

Experiment 2

Figure 2

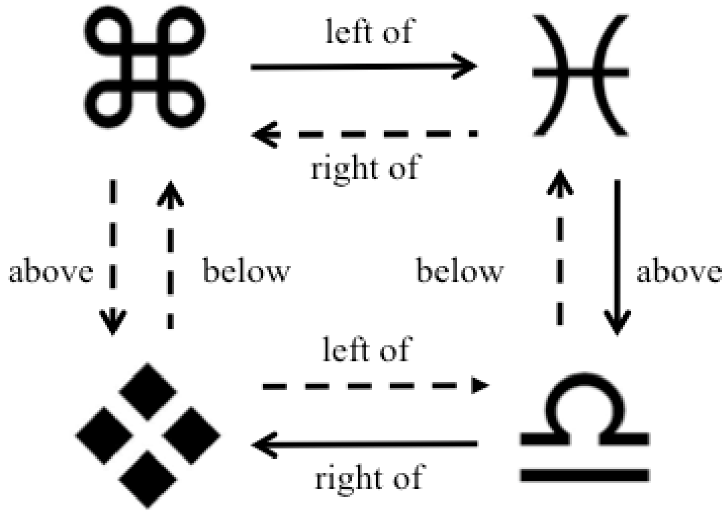


Figure 3

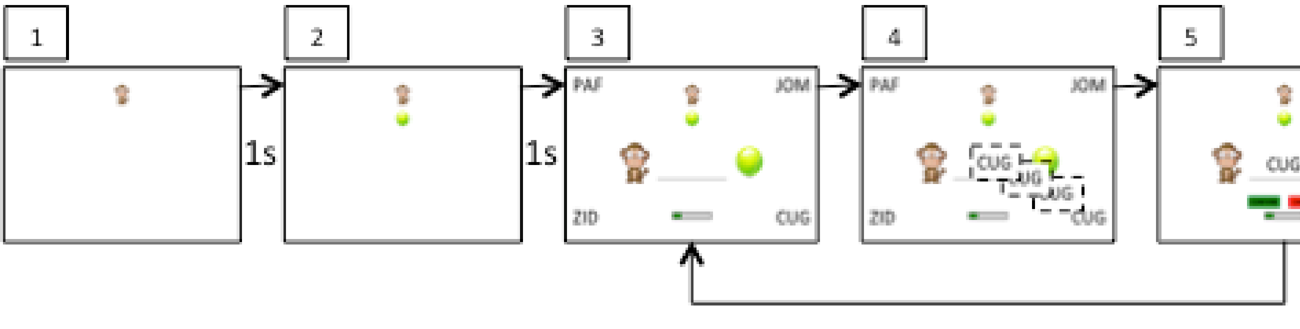


Figure 4

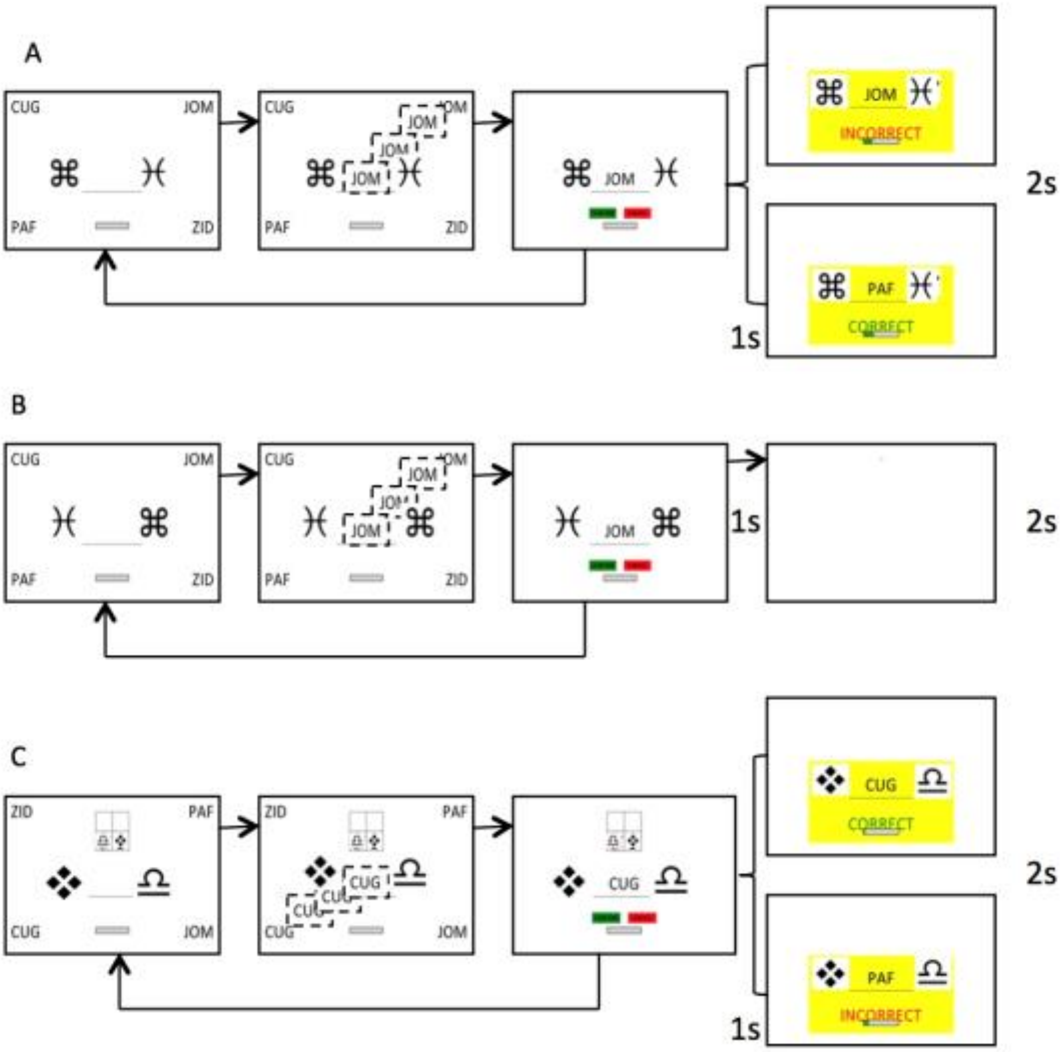
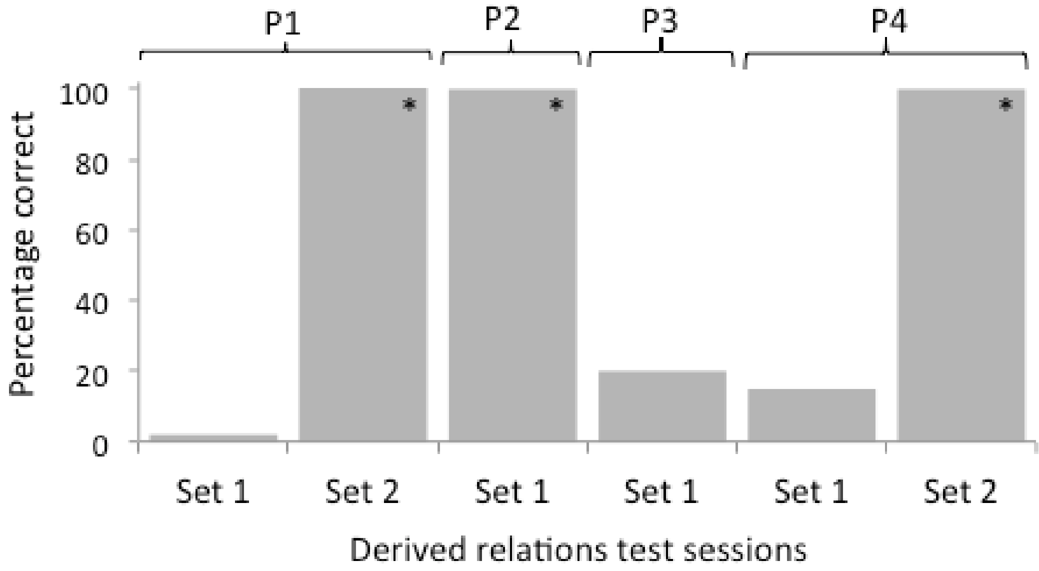
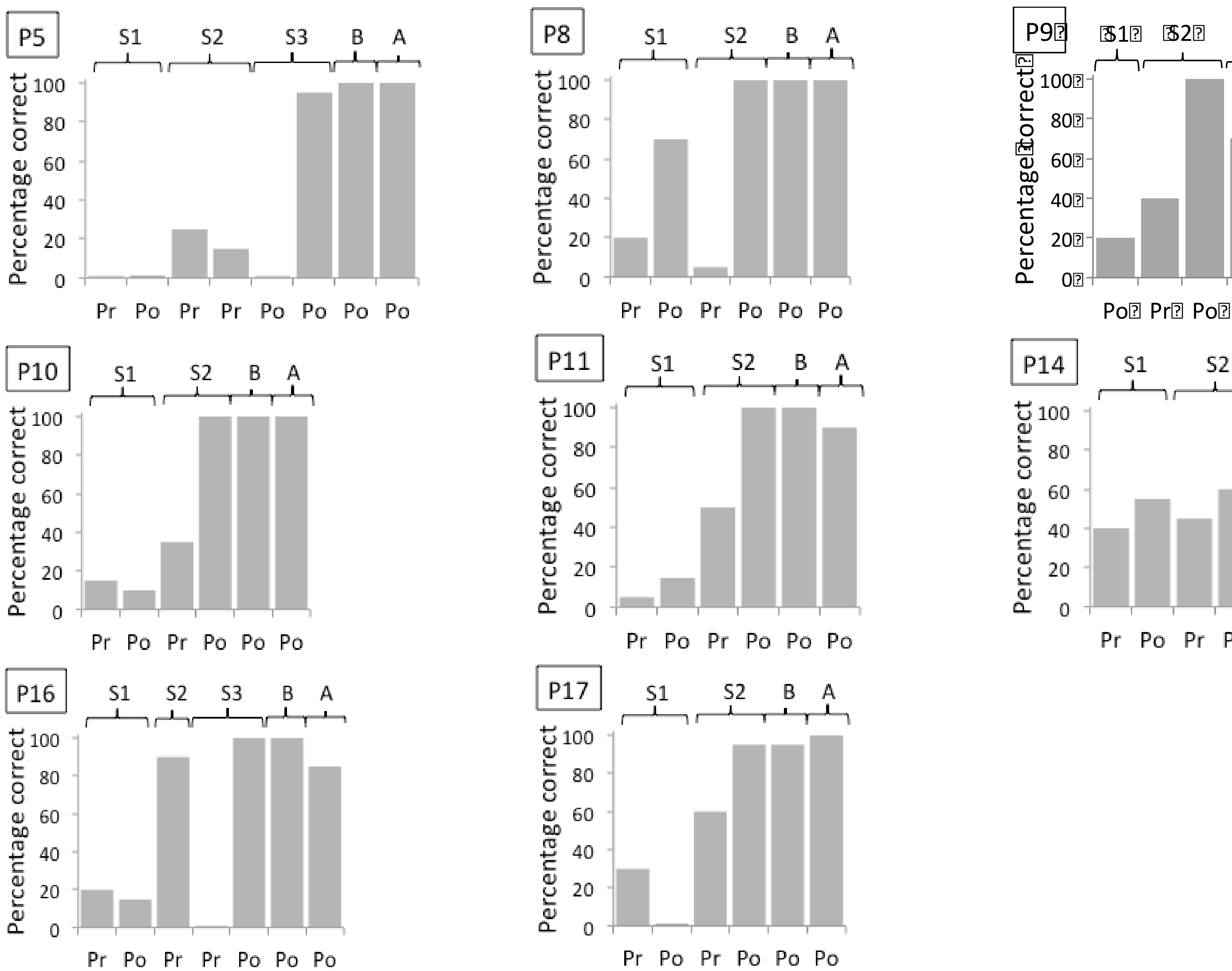


Figure 5


P1
P2
P3
P4
100
80
60
40
20
0
Percentage correct
*
*
*
Set 1
Set 2
Set 1
Set 1
Set 1
Set 2
Derived relations test sessions



P5
S1
S2
S3
B
A
Percentage correct
0
20
40
60
80
100
Pr Po Pr Pr Po Po Po Po
P8
S1
S2
B
A
Percentage correct
Pr Po Pr Po Po Po
P9
S1
S2
Percentage correct
Po Pr Po
P10
S1
S2
B
A
Percentage correct
Pr Po Pr Po Po Po
P11
S1
S2
B
A
Percentage correct
Pr Po Pr Po Po Po
P14
S1
S2
Percentage correct
Pr Po Pr
P16
S1
S2
S3
B
A
Percentage correct
Pr Po Pr Pr Po Po Po
P17
S1
S2
B
A
Percentage correct
Pr Po Pr Po Po Po

Table 1

Trained and tested relations during arbitrary phases (Set 1)

Type	Relation (Stimuli)	Relation type	Sample 1	Sample 2
Arbitrary	A left of B	Trained	A1	B1
Arbitrary	B above C	Trained	B1	C1
Arbitrary	C right of D	Trained	C1	D1
Arbitrary	B right of A	Derived (ME)	B1	A1
Arbitrary	D left of C	Derived (ME)	D1	C1
Arbitrary	C below B	Derived (ME)	C1	B1
Arbitrary	A above D	Derived (CE)	A1	D1
Arbitrary	D below A	Derived (CE)	D1	A1

Note. Letters denote the stimulus position in the network (A= top left, B= top right, C= bottom right, and D=bottom left) and the numeral denotes the set number; "ME" = mutually entailed relation and "CE" = combinatorial entailed relation.

Table 2

Trial configurations (sample, contextual cue) for nonarbitrary training and testing in Experiment 1 (Phase 1), arbitrary training and testing (Phases 2-3), and remedial training (Phase 4).

Phase	Task	Sr+	Crit.	Phase	Task	Sr+	Crit.
1a	S1/S2 [left of]	100%	7/8	2	A1/B1 [left of]	100%	12/12
					B1/C1 [above]		
1a	S1/S2 [above]	100%	7/8		C1/D1 [right of]		
1a	S1/S2 [left of]	100%	7/8	3	*B1/A1 [right of]*	0%	18/20
	S1/S2 [above]				*C1/B1 [below]*	*derived*	*derived*
					D1/C1 [left of]	100%	11/12
1a	S1/S2 [below]	100%	7/8		*D1/A1 [below]*	trained	trained
					A1/D1 [above]		
1a	S1/S2 [left of]	100%	11/12		A1/B1 [left of]		
	S1/S2 [above]				B1/C1 [above]		
	S1/S2 [below]				C1/D1 [right of]		
1a	S1/S2 [right of]	100%	7/8	4	*B1/A1 [right of]*	100%	18/20
					C1/B1 [below]		*derived*
1a	S1/S2 [left of]	100%	7/8		*D1/C1 [left of]*		11/12
	S1/S2 [above]				*D1/A1 [below]*		trained
	S1/S2 [below]				*A1/D1 [above]*		
	S1/S2 [right of]				A1/B1 [left of]		
					B1/C1 [above]		
1b	S3/S4 [left of]	100%	7/8		C1/D1 [right of]		
	S3/S4 [above]						
	S3/S4 [below]						
	S3/S4 [right of]						
	S4/S3 [left of]						
	S4/S3 [above]						
	S4/S3 [below						
	S4/S3 [right of]						
1c	S5/S6 [left of]	0%	7/8				
(test)	S5/S6 [above]						
	S5/S6 [below]						
	S5/S6 [right of]						
	S6/S5 [left of]						
	S6/S5 [above]						
	S6/S5 [below]						
	S6/S5 [right of]						

Note. The trials for arbitrary and nonarbitrary relational training and testing are described using the following convention: The first sample stimulus is described in capitals, followed by the second sample stimulus, followed by the correct comparison stimulus (contextual cue), which is shown in brackets. For example, the notation B1/C1 [*above*] indicates that in the presence of the first sample stimulus B1, followed by C1, dragging and dropping the contextual cue *Above* was reinforced whereas, selecting *Right of, Left of,* or *Below* was not.

Table 3

Total number of trials to criterion during each phase of Experiment 1.

	Phases						
P	1a	1b	1c	Set #	2	3	4
1	68	8	16	1	36	*32*	64*
				2	48	32*	
2	68	8	8	1	24	32*	
3	76	72	8	1	24	*32*	*192*
4	76	16	8	1	36	*32*	128*
				2	24	32*	

Note. An asterisk in testing indicates criterion level responding during a test for derived relations. Failure to meet pass criteria during a test is highlighted in italics.

Table 4

Trial configurations for all Phases in Experiment 2

Phase	Task	Sr^+	Crit.	Phase	Task	Sr^+	Crit.
1a	S1/S2 [left of]	100%	7/8	2	*B1/A1 [right of]*	0%	18/20
					C1/B1 [below]		
1a	S1/S2 [above]	100%	7/8		*D1/C1 [left of]*		
					D1/A1 [below]		
1a	S1/S2 [left of]	100%	7/8		*A1/D1[above]*		
	S1/S2 [above]						
				3	A1/B1 [left of]	100%	12/12
1a	S1/S2 [below]	100%	7/8		B1/C1 [above]		arbitrary
					C1/D1 [right of]		7/8
1a	S1/S2 [left of]	100%	11/12		S5/S6 [left]		non-arb
	S1/S2 [above]				S5/S6 [right]		
	S1/S2 [below]				S5/S6 [above]		
					S5/S6 [below]		
1a	S1/S2 [right of]	100%	7/8				
				4	*B1/A1 [right of]*	0%	18/20
1a	S1/S2 [left of]	100%	7/8		*C1/B1 [below]*	*derived*	*derived*
	S1/S2 [above]				*D1/C1 [left of]*	100%	11/12
	S1/S2 [below]				*D1/A1 [below]*	trained	trained
	S1/S2 [right of]				*A1/D1[above]*		
					A1/B1 [left of]		
1b	S3/S4 [left of]	100%	7/8		B1/C1 [above]		
	S3/S4 [above]				C1/D1 [right of]		
	S3/S4 [below]						
	S3/S4 [right of]			5	*B1/A1 [right of]*	100%	18/20
					C1/B1 [below]		*derived*
1c	S4/S3 [above]	100%	7/8		*D1/C1 [left of]*		11/12
	S3/S4 [below]				*D1/A1 [below]*		trained
					A1/D1[above]		
1c	S4/S3 [above]	100%	7/8		A1/B1 [left of]		
	S3/S4 [below]				B1/C1 [above]		
	S4/S3 [below]				C1/D1 [right of]		
	S3/S4 [above]						
				6a	D1/B1 [left of]	100%	12/12
1c	S4/S3 [right]	100%	7/8		B1/A1 [above]		arbitrary
	S3/S4 [left]				A1/C1 [right of]		7/8 non-
					S5/S6 [left]		arb
1c	S4/S3 [right]	100%	7/8		S5/S6 [right]		
	S3/S4 [left]				S5/S6 [above]		
	S4/S3 [left]				S5/S6 [below]		
	S3/S4 [right of]						

1c	S3/S4 [left of]	100%	7/8	6b	*B1/D1 [right of]*	0%	18/20
	S3/S4 [above]				*A1/B1 [below]*	*derived*	*derived*
	S3/S4 [below]				*C1/A1 [left of]*	100%	11/12
	S3/S4 [right of]				*C1/D1 [below]*	trained	trained
	S4/S3 [left of]				*D1/C1 [above]*		
	S4/S3 [above]				D1/B1 [left of]		
	S4/S3 [below				B1/A1 [above]		
	S4/S3 [right of]				A1/C1 [right of]		
1d	S5/S6 [left of]	0%	7/8	7a	*Trial types*		
(test)	S5/S6 [above]				*identical to*		
	S5/S6 [below]				*Phase 3*		
	S5/S6 [right of]						
	S6/S5 [left of]			7b	*Trial types*		
	S6/S5 [above]				*identical to*		
	S6/S5 [below				*Phase 4*		
	S6/S5 [right of]						

Table 5

Total number of trials to criterion and test scores during Experiment 2

Participant	Phases									
	1a	1b	1c	1d	Set	2	3	4	5	6a
5	68	8	64	16	1	0/20	40	*32*	64	
					2	5/20	40	*32*	64	
					3	0/20	60	32*		40
6	68	8	56	8	1	8/20	*120*			
7	68	8	32	8	1	3/20	*120*			
8	78	8	56	8	1	4/20	80	*64*	32	
					2	1/20	40	32*		60
9	68	8	56	16	1	-	160	*32*	96	
					2	6/20	60	32*		60
10	104	8	64	16	1	3/20	120	*32*	96	
					2	7/20	120	96*		60
11	48	8	*80*							
	32[b]	8[b]	40[b]	8	1	1/20	60	*32*	192 64[c]	
					2	10/20	40	32*		40
12	92	40	168	8	1	11/20	*120*			
13	60	16	72	8	1	5/20	*240*			

14	76	8	48	8	1	8/20	100	*32*	128	
					2	9/20	40	*96*	64	
					3	2/20	40	*32*	64	
					4	8/20	40	*32*	64	
					5	0/20	60	*32*		
15	156	64	*128*							
			96[b]							
16	68	32	40	8	1	4/20	120	*32*	64	
					2	18/20				
					3	0/20	40	32*		40
17	56	8	40	8	1	6/20	60	*64*	64	
					2	12/20	40	32*		40

Note. An asterisk in testing indicates criterion level responding during a test for derived relations. Failure to me or testing is highlighted in italics; [b] indicates training sessions involving adapted nonarbitrary training; [c] indicates traini remedial training. See text for further details.

Figure Captions

Figure 1. Stimuli used during nonarbitrary and arbitrary training and testing in Experiments 1 and 2. *Wingdings* were employed within stimulus sets 1-3 for Participants 1-7. *Fribbles* were employed within stimulus sets 1-5 for Participants 8-17. See text for details.

Figure 2. Schematic diagram of the trained (solid lines) and tested (dashed lines) spatial relational network. See text for details.

Figure 3. Schematic diagram of the sequence and format of stimulus presentation during nonarbitrary relational training. These illustrative stimuli are drawn from the sets and relations shown in Figure 1. In Panel 1, S1 is presented in isolation at the center, top of the screen. In Panel 2, S2 appears directly to the right of S1. In Panel 3, all four relational cues are presented in each corner of the display at the same time as the sample stimulus S1____S2. In Panel 4, the participant selects either the *Left of*, *Right of*, *Above*, or *Below* from the relational cues by dragging and dropping to "complete a statement". Panel 5 depicts the confirmation screen in which the 'confirm' (green) and 'cancel' (red) buttons are presented. Finally, either correct or incorrect feedback is presented (right hand panels).

Figure 4. Schematic diagram of the sequence and format of stimulus presentation during arbitrary relational training (A), arbitrary relational testing (B), and remedial training (C) in both experiments. Panel A shows the sequence of screen displays and response options presented during arbitrary training of a "Left of" relation. Panel B shows the sequence of screen displays and response options presented while testing an arbitrary *Right of* relations. Panel C depicts the nonarbitrary remedial training trials only (i.e., A1/B1 [left of], B1/C1 [above], and C1/D1 [right of]). During these trials, participants were provided with additional onscreen stimuli. In addition to the sample stimuli '1A____1B', a 2 x 2 grid appeared which contained duplicates of 1A and 1B in the respective (nonarbitrary) positions in the relational network. All other trials during remedial training were presented in the same fashion as that shown in Panel A (i.e., including feedback). See text for details.

Figure 5. Percentage correct scores for derived relations during posttest sessions during Experiment 1. An asterisk indicates criterion level responding for derived relations.

Figure 6. Percentage correct scores for derived relations during Phase 2 (pretest), Phase 4 (posttest), Phase 6 (reassignment test), and Phase 7 (reversal test) during Experiment 2. Note: Pr= Pretest, Po= Posttest, S1= Stimulus set 1, S2= Stimulus set 2, S3= Stimulus set 3, S4= Stimulus set 4, S5= Stimulus set 5, B= Reassignment test, A= Reversal test.