JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2009, **91,** 105–126 NUMBER 1 (JANUARY)

PRODUCING AND RECOGNIZING ANALOGICAL RELATIONS

Regina Lipkens

PSYCHIATRIC HOSPITAL SANCTA MARIA
SINT-TRUIDEN, BELGIUM

AND

Steven C. Hayes

UNIVERSITY OF NEVADA, RENO

Analogical reasoning is an important component of intelligent behavior, and a key test of any approach to human language and cognition. Only a limited amount of empirical work has been conducted from a behavior analytic point of view, most of that within Relational Frame Theory (RFT), which views analogy as a matter of deriving relations among relations. The present series of four studies expands previous work by exploring the applicability of this model of analogy to topography-based rather than merely selection-based responses and by extending the work into additional relations, including nonsymmetrical ones. In each of the four studies participants pretrained in contextual control over nonarbitrary stimulus relations of sameness and opposition, or of sameness, smaller than, and larger than, learned arbitrary stimulus relations in the presence of these relational cues and derived analogies involving directly trained relations and derived relations of mutual and combinatorial entailment, measured using a variety of productive and selection-based measures. In Experiment 1 participants successfully recognized analogies among stimulus networks containing same and opposite relations; in Experiment 2 analogy was successfully used to extend derived relations to pairs of novel stimuli; in Experiment 3 the procedure used in Experiment 1 was extended to nonsymmetrical comparative relations; in Experiment 4 the procedure used in Experiment 2 was extended to nonsymmetrical comparative relations. Although not every participant showed the effects predicted, overall the procedures occasioned relational responses consistent with an RFT account that have not yet been demonstrated in a behavior-analytic laboratory setting, including productive responding on the basis of analogies.

Key words: analogical relations, derived stimulus relations, language, Relational Frame Theory, mutual entailment, combinatorial entailment, matching to sample, productive equivalence, humans

---

The derivation and use of analogies and metaphors is one of the most productive aspects of human responding. Sets of skills and domains of knowledge are quickly brought to bear on new areas via analogy and metaphor. Skills in analogical reasoning are often used as a metric of intelligent behavior (Sternberg, 1977a) and to predict academic success (e.g., the use of analogy problems in the Graduate Record Examination or the Miller Analogies Test). The effective use and promotion of analogical abilities is important in many areas of education (Kolodner, 1997). For example, human problem solving has been shown to be improved by facilitating the use of analogy, including when compared to direct instructions (Brown, 1989; Kamouri, Kamouri, & Smith, 1986; Wong, 1993).

Analogy is important not just because it is a behavior of known applied importance, but also because it seems key to an understanding of human language and cognition itself (Eysenck & Keane, 1995; Ortony, 1993). Whereas analogical reasoning has received considerable attention from cognitive scientists (e.g., Getner & Holyoak, 1997; Holyoak & Thagard, 1995, 1997), the accounts are generally structuralistic and the precise nature and history of analogy is often unclear. This may be changing, as some cognitive scientists have begun to conclude that the extension of relational learning depends on analogies learned via multiple exemplars (Doumas, Hummel, & Sandhofer, 2008), arriving at a

---

This study was done in partial fulfillment of the requirements for a doctoral degree granted to Regina Lipkens by the University of Nevada.

Correspondence and requests for reprints should be addressed to Steven C. Hayes, Department of Psychology, University of Nevada, Reno, NV 89557-0062 (e-mail: stevenchayes@gmail.com).

doi: 10.1901/jeab.2009.91-105

position somewhat similar to the behavioral account known as Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001).

Behavior analysts have historically emphasized the role of abstraction in analogy and metaphor (Skinner, 1957) but have done so interpretively. For Skinner, analogy is a process of abstraction, via an extended tact, of common physical properties (e.g., 1957, p. 93). This is a reasonable starting point but it leaves untouched purely verbal analogies (Stewart, Barnes-Holmes, & Roche, 2004), and the histories required for such performances. Furthermore, no empirical literature has emerged from these interpretations. In behavior analysis, the empirical analysis of analogy has occurred in the context of RFT (e.g., Barnes, Hegarty, & Smeets, 1997; Barnes-Holmes et al., 2005; Lipkens, 1992; Stewart et al., 2004).

Analogy is a broad concept that comes from the Greek *ana logon*: ''according to a ratio'' (*New Encyclopædia Britannica, Micropædia,* Vol. I, 1987, p. 367). Originally, the Greek mathematicians used the word 'analogia' to denote a similarity in proportional relationships (e.g., Euclid, trans.1956, p. 112–115). Another form of analogy noted by the Greeks is that of inferring similarity of function (*New Encyclopædia Britannica, Micropædia,* Vol. I, 1987, p. 367). In this sense, analogy designs a direct linking of two terms (the analogates) that are compared with respect to a property attributed in some way to both analogates. In this form of analogy one of the terms must always be the ''prime analogate'', on which the analogous property is predicated formally or intrinsically, while the other receives an attribution in a secondary sense by virtue of some relevant, real relation to the prime analogate (*Encyclopedia of Philosophy,* Vol. I, 1967, p. 95).

These two forms of analogy are known respectively as analogy of proportionality and analogy of attribution (*Encyclopedia of Philosophy,* Vol. I, 1967, p. 95; *New Encyclopædia Britannica, Micropædia,* Vol. I, 1987, p. 367). In the *Metaphysics* (Bk, IX, Ch. 6, 1048b) Aristotle (trans. 1941b, p. 826) stated the formulas of the two kinds of analogy: ''As A is in B or to B, C is in D or to D'' or in another translation ''The way that this is in or is related to this is like the way that that is in or is related to that'' (Aristotle, trans. 1963, p. 110). Aristotle also formulated the classic proportion schema: A:B::C:D, then (alternando) A:C::B:D, and therefore (componendo) A+C:B+D::A:B (Aristotle, trans 1941a, p. 1007).

The behavior-analytic empirical work in the area originally emphasized equivalence or frames of coordination over other forms of relational responding. Indeed, Barnes et al. (1997) originally defined analogy as a matter of ''equivalence–equivalence relations'' in a study showing that participants would relate equivalence relations to other equivalence relations and nonequivalence relations to other nonequivalence relations. Over time it has become clear that the core relational performance can be thought of more generally. In line with the distinction between analogy of proportionality and analogy of attribution, Stewart, Barnes-Holmes, Roche, and Smeets (2001) showed that derived relations could be based on the abstraction of formal properties that are shared among events (what relational frame theorists term ''pragmatic verbal analysis,'' Hayes, Gifford, Townsend, & Barnes-Holmes, 2001). This begins to provide a process account for a core feature of Skinner's ideas on analogy. Stewart et al. (2004) and Barnes-Holmes et al. (2005) extended the empirical account to relations of difference, showed that novel relations could emerge using procedures other than matching-to-sample, and found that these differences were reflected in neurobiological measures. As equivalence itself has begun to be analyzed experimentally as a relational operant (e.g., Luciano, Becerra, & Valverde, 2007) in the hands of relational frame theorists, the behavioral core of analogy no longer seems to be a matter of equivalence, per se, but of relating relations, both trained and derived (Barnes-Holmes et al., 2005).

Within RFT, relating relations is particularly important because it arguably explains in part the degree of generativity of language and how language can lead to the creation of precise but abstract concepts (Stewart, Barnes-Holmes, Hayes, & Lipkens, 2001). Rather than bringing items one by one into networks of relational responses, relating relations allows entire sets of such responses to impact other sets, modeling some of the generativity seen in human language and cognition. This could be why most forms of complex reasoning, such as scientific and mathematical skills, are based in part on reasoning by analogy (e.g., Clement, 1988; Novick & Holyoak, 1991; Polya, 1954). The refinement of stimulus control this process affords can also produce new forms of behavior with subtle features that seem to go beyond simple instructional control. A person in psychological difficulty reading about the ''bus of life'' may see the source and cost of their problems in new ways. They may realize they have been ''putting life on hold while fighting a war within'' much as a bus driver has to pull over to go back and try to throw off passengers. Just explaining the full implications of such an insight could take several times longer than the time it took to create them via analogy and metaphor. Indeed, as Skinner (1989) noted, most abstract concepts in language are frozen analogies and metaphors, suggesting a central role for this process in the creation of abstract concepts of any kind.

The present set of experiments advances the behavior-analytic empirical literature on analogy in several ways. If derived stimulus relations are a key feature of an adequate model of productive human language, researchers need to go beyond a focus on stimulus selection as a measure of these relations. The derived stimulus relation literature contains a few examples of the use of productive response measures (e.g., Lipkens, Hayes, & Hayes, 1993) but only a few. Several studies have shown that selection-based measures can reveal a derived relation whereas a response topography-based measure does not (e.g., Polson, Grabavac, & Parsons, 1997; Polson & Parsons, 2000). Extending the literature into this area is important because in natural situations response forms are a key feature of analogous performances. Intelligence tests ask respondents to fill in blanks with a word (e.g., bird is to nest as a spider is to a _____); in problem-solving, analogies suggest novel actions, not merely selecting options from an array. Evidence from the cognitive literature also shows that in composing an analogy, requiring a more elaborated response form such as writing improves the ability to apply the analogy (Klein, Piacente-Cimini, & Williams, 2007). To date, no behavioral studies have shown that analogical relations can lead to the regulation of unique response forms.

Performances on recognition tests are generally better than those on recall tests (e.g., Nelson, McKinney, Gee, & Janczura, 1998), and in a similar fashion producing relata may be somewhat more difficult than selecting them. From an RFT point of view, however, if participants can say what they hear or write what they see, this difficulty should not be insurmountable provided that the task properly emphasizes derived relational responding over rote learning of conditional discriminations. In the present study several features occurred concurrently with topography-based response tasks (e.g., selecting relata, selecting relations, the use of multiple relations) so that a focus would remain on derived relational responding.

It is also important to distinguish between the construction of an analogy (that is, the derivation of a relation among relations, Stewart, Barnes-Holmes, Hayes, & Lipkens, 2001) and its application to entirely new events so as to expand a relational network. This distinction is similar to the traditional distinction between the eduction of correlates and of the eduction of relations (Spearman, 1973, Ch. 5-7), but with some exceptions (Stewart et al., 2004) the behavioral literature to date has largely emphasized the ability to relate relations, not the use of that process to give relational functions to novel events. Both features of analogy are important, and the present experiments give additional attention to that issue.

Finally, it is theoretically important to expand the set of relational frames empirically examined in research on analogy, to relations that are difficult or impossible to address via conditional control over stimulus equivalence. Some attempts have been made to think of frames of opposition and difference as matters of equivalence class control by negative stimuli (Johnson & Sidman, 1993), but so far as we know no one has yet extended equivalence to handle nonsymmetrical relations such as temporal or comparative relations (Hayes & Barnes, 1997). Sidman (2008) is explicit that he is not attempting to do so, stating that his analysis is meant as ''a *limited* theory in that it does not cover other kinds of relations than equivalence, as for example, relational frame theory attempts to do'' (p. 331, emphasis in original).

Comparative relations are particularly important to examine in this light because there is experimental evidence that comparative relations can be trained as relational operants in participants who do not have them (Berens & Hayes, 2007). For these reasons, comparative relations were included in the relational networks that were the basis of analogies in the present study, along with same, opposite and different.

EXPERIMENT 1: FINDING ANALOGIES WITH SAME/OPPOSITE RELATIONS

The question asked in Experiment 1 is whether a participant will be able to find an analogy between a relation in one network and a relation in another network based on same and opposite relations, even when different trained relations underlie the analogy, but only the derived relations are analogous. This is similar to what has been shown previously (e.g., Stewart et al., 2004) but in previous work same and different relations have been assessed and no topography-based measures were included. A common sense example may be how cat owners may derive an analogy such as ''cats are dictators'' based on their knowledge of both domains.

Method

Participants

Seventeen participants started SAME / OPPOSITE pretraining. Eleven of them completed the three components of pretraining (described below) that were needed to establish specific relational cues, demonstrate their arbitrary applicability, and teach participants to treat the mere presentation of stimuli pairs as noting that they were related. Four of these SAME / OPPOSITE pretrained participants served in the present experiment—one withdrew after the second session and thus has only partial results. The remaining 7 served in Experiment 2. All participants were students in introductory psychology who participated for bonus credits.

Setting and Apparatus

During pretraining the participant and the experimenter were seated in a small room at a table opposite each other. The stimuli used as sample and comparison stimuli were black and white and colored figures copied from magazines or drawn by the experimenter. The stimuli used to indicate the kind of relation were black and white novel geometrical figures. During all other phases of the experiments the participant was seated in a small room before a 14'' computer screen and a keyboard.

Pretraining

The basic relational skills assessed in pretraining have been shown in many previous RFT studies (Hayes, Barnes-Holmes, et al., 2001). Because the primary focus in this series of experiments is whether such skills can be brought to bear on relations among relations, the 6 participants (35% of the total) who failed repeatedly in pretraining (their relational cue pretraining exceeded 1 hr; they had three unsuccessful attempts to complete each phase of the pretraining) were dismissed.

Relational cue pretraining. The purpose of relational cue pretraining was to bring nonarbitrary stimulus relations (that is, relations based on the formal properties of the relata) of sameness and oppositeness under the control of arbitrary contextual cues that could later be used to establish arbitrary stimulus relations. At the start of the first session the following instructions were given (no other instructions were given in this part of pretraining):

> During the first phase of this experiment I will present several stimuli on cards. After this phase, which will last between 20 and 50 minutes, the experiment will be continued before a computer. During this phase I will present a stimulus at the top, one at the center, and two or three stimuli at the bottom of this table. You have to select one of the bottom stimuli. In the beginning I will tell you whether your choice is correct or incorrect. On the basis of my feedback you have to figure out which stimulus you have to select. Later on I will give no feedback: so then I will not tell you whether your choice was right or wrong.

The stimuli used as sample and comparison stimuli were black-and-white and colored figures copied from magazines and pasted on orange paper. The stimuli used to indicate the kind of relation were novel black and white geometrical figures on green paper. Two different geometrical stimuli were used as contextual stimuli identifying a specific relation. We will refer to these second-order conditional stimuli by the generic name of ''relational cues,'' and when speaking of a specific type we will refer to them in capitals as the SAME or OPPOSITE stimuli.

For reasons that will become more apparent in Experiments 3 and 4, some trials used two comparisons, whereas others used three. The four stimulus sets from which the three comparison problems were drawn had two subsets of three stimuli as follows: two subsets of drawings of a small, medium, and large circle; drawings of 3, 6, and 28 stars, and of 3, 6, and 28 dots; pictures of 1, 3, and 9 airplanes and of 1, 3, and 9 persons; and two subsets consisting of different pictures of a baby, an adolescent, and an elderly man. The four stimulus sets from which the two comparison problems were drawn had two pairs of stimuli that characterized the same formal relation but in a slightly different way: a picture of a stove and of a child eating ice cream, and a picture of an open fire and of two penguins; a picture of a motorcyclist and of two elderly people carrying heavy bags walking on a beach, and a picture of a lion and of a turtle; a picture of a laughing girl and of a sad man, and a picture of a laughing man and of a crying child; a picture of a fat rat and a schematic drawing of a thin person, and a picture of a thin lady and of a fat man.

During trials from three comparison stimulus sets, one of the extreme stimuli from one subset was used as a sample (e.g., one or nine airplanes) and the other subset as comparisons (e.g., one, three, or nine persons). During trials from the two comparison stimulus sets, one of the stimuli from a paired subset served as a sample (e.g., a turtle) for the two stimuli from the other pair (e.g., a motorcyclist and two elderly people carrying heavy bags walking on a beach).

The sample stimulus appeared at the top of the table, the relational stimulus at the center, and the comparison stimuli in a row in random order at the bottom of the table. During the training when feedback was given, the experimenter said ‘‘correct’’ or ‘‘yes’’ when the response was correct and ‘‘wrong’’ or ‘‘no’’ when the choice was incorrect. For each stimulus set, participants were taught that in the presence of one geometrical stimulus (the SAME stimulus) selection of the comparison that was similar to the sample was correct; in the presence of the second geometrical stimulus (the OPPOSITE stimulus), selection of the comparison that was opposite to the sample was correct. When four successive correct answers were given, with feedback, the trials were presented in random order without feedback. If performance was still 100% accurate, a new stimulus set was presented; in the no feedback phase multiple stimulus sets were mixed. The final criterion for completion of the relational cue pretraining phase was errorless performance during the presentation of four consecutive novel sets of stimuli without feedback. Failures led to retraining.

Arbitrary relation pretraining. There were two purposes of this phase: 1) to see if participants would derive same and opposite relations among arbitrary stimuli based on the presence of the SAME and OPPOSITE relational cues established in relational cue pretraining, and 2) to establish three different kinds of derived relational performance: selecting the correct relata, producing the correct relata, and selecting the correct relational cue.

During this and all subsequent phases of the experiment the participant was seated before a computer screen and a keyboard. The stimuli used as sample and comparison stimuli were drawn from a list developed by Underwood and Schultz (1960, p. 308-428) of two- or three-letter stimuli with a low frequency of occurrence in the English language and (for two-letter stimuli) a low response frequency to single letters. The relational cues were those from training in the previous phase but now presented on the computer monitor. All stimuli used in this phase and in the experimental phases of all experiments are shown in Figure 1.

In arbitrary relation pretraining arbitrary stimulus relations of sameness and oppositeness were shown to be under the contextual control of the SAME and OPPOSITE relational cues. Two sets of three stimuli were used as sample and comparison stimuli: (1) DOX (which we will call ‘‘X1’’ for ease of methodological description), BAF (‘‘Y1’’), and LIQ (‘‘Z1’’); and (2) JUN (‘‘X2’’), GEK (‘‘Y2’’), and PYM (‘‘Z2’’). During feedback trials, the word ‘‘right’’ appeared in the left upper corner of the screen when the response was correct and the word ‘‘wrong’’ appeared together with a low-pitched sound when the response was incorrect.

Participants were taught the following relations: X1 is the same as Y1, X1 is the opposite

Words Used As Sample and Comparison Stimuli in Arbitrary Relation Pretraining

(X1) DOX	(X2) JUN
(Y1) BAF	(Y2) GEK
(Z1) LIQ	(Z2) PYM

Words Used As Sample and Comparison Stimuli in Experiment 1 and 2

(A1) AK	(A2) EW	(A3) IH	(A4) OC
(B1) AV	(B2) EJ	(B3) IZ	(B4) OX
(C1) AF	(C2) EQ	(C3) IP	(C4) OG

Words Used As Sample and Comparison Stimuli in Experiment 3 and 4

(A1) AK	(A2) EW
(B1) AV	(B2) EJ
(C1) AF	(C2) EQ

Geometrical Stimuli Used As Relational Cues in Pretraining, Training and Tests in Experiment 1 and 2

(SAME) (OPPOSITE)

Geometrical Stimuli Used As Relational Cues in Pretraining, Training and Tests in Experiment 3 and 4

(SAME) (SMALLER) (LARGER)

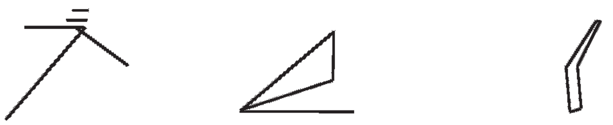

Fig. 1. Stimuli used in all four experiments.

of Z1, X2 is the same as Y2, and X2 is the opposite of Z2. Because one purpose of this phase was to establish different forms of relational performance, these relations were trained in three ways (for clarity of presentation, we will adopt the convention throughout of capitalizing the type of relational task). Figure 2 presents screen examples of the three different tasks. The SELECTING RELATA task was similar to that used in the previous phase: The sample stimulus appeared at the top of the screen, the relational stimulus at the center, the comparison stimuli in a row at the bottom, and the participant had to select one of the comparisons. In the PRODUCING RELATA task, the sample stimulus appeared at the top of the screen, the relational stimulus at the center, and three dots and a question mark (...?) were at the bottom. The participant had to produce the correct comparison stimulus by typing the letters on the keyboard. In the SELECTING RELATION task, a complex sample stimulus consisting of the sample stimulus, three dots and a question mark and the correct comparison stimulus (e.g., X1 ...? Y1) appeared at the top and the two relational cues (SAME and OPPOSITE) at the bottom. The participant had to select one of the relational cues. These formal differences were described in the final instructions given to participants:

> Two kinds of problems will be presented to you. In the first problem, first one or more stimuli will appear at the top of the screen and one stimulus will appear at the center of the screen. Then, two, three, or four stimuli will appear in a row at the bottom of the screen. You have to choose one of these stimuli. In case there are two stimuli at the bottom, you make your selection by pressing key 1 or 0 located at the top of your keyboard for respectively the left or the right stimulus. In case there are three stimuli at the bottom, you make your selection by pressing key 1, 5, or 0 for respectively the left, middle, or right stimulus. In case there are four stimuli at the bottom, you press key 1, 4, 7, or 0 for

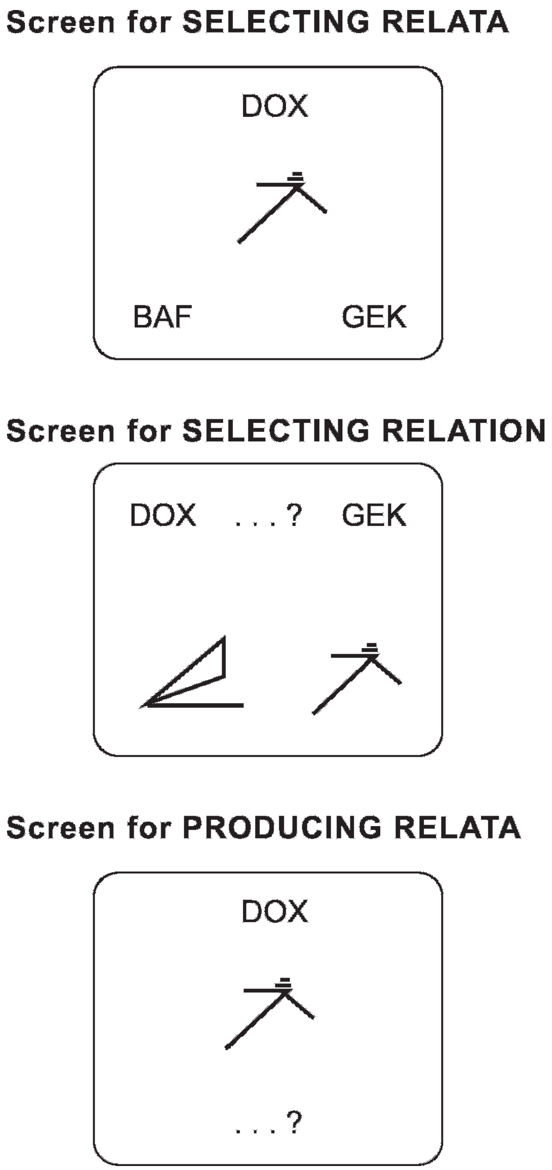

Fig. 2. Examples of the screen arrangements for the three different arbitrarily relation pretraining tasks in Experiments 1 and 2. Stimuli can be linked to sets and relations using Figure 1. In the example of the SELECTING RELATA screen, X1 is the sample, and the SAME relational cue sets the occasion for selecting either Y1 or Y2. In the SELECTING RELATION screen, X1 ___? Y1 is the sample and the participant must select either the OPPOSITE or SAME relational cue. In the PRODUCING RELATA example, X1 is the sample, and in the presence of the SAME relational cue, relata must be typed in.

respectively the extreme left, middle left, middle right, or extreme right stimulus. In the second problem, first, one or more stimuli will appear at the top and one stimulus will appear at the center of the screen, just as in the first problem. Then at the bottom three dots with a question mark (...?) will appear. Instead of selecting a stimulus as in the first problem, you have to type now some letters using the keys of the keyboard to produce a stimulus you think is the correct one for that problem. If you made an error you can delete letters by pressing the key with the arrow. After typing in some letters you have to press the enter key to continue the experiment. You cannot write anything down during the experiment.

Participants first learned that X1 is the same as Y1 and X1 is the opposite of Z1 using the three different tasks. Participants were then taught in the same manner that X2 is the same as Y2 and X2 is the opposite of Z2. Complete accuracy in a four-trial block with a stimulus set on a particular relational task (e.g., SELECTING RELATA) led to the next relational task, and when all three had been trained, to the next stimulus set. Then both stimulus sets were presented in mixed order in four trial blocks, first with feedback and then without. After the participant responded 100% accurately in each relational task in the mixed stimulus set, the test phase started. Using all three types of tasks (SELECTING RELATA, SELECTING RELATION, and PRODUCING RELATA), the participants were tested for the derived relations of mutual entailment (Y1 is the same as X1, Z1 is the opposite of X1, Y2 is the same as X2, and Z2 is the opposite of X2), and of combinatorial entailment (Y1 is the opposite of Z1, Z1 is the opposite of Y1, Y2 is the opposite of Z2, and Z2 is the opposite of Y2). Feedback was never given during test trials. The criterion for a successful arbitrary relation pretraining was errorless performance in each relational task for both mutual and combinatorial entailment in four trial blocks (six tests in total; 24 out of 24 trials correct). Failures led to retraining.

Stimulus pair pretraining. The purpose of this phase of pretraining was to establish the meaning of a particular form of presentation of stimulus pairs: that ''X–Y'' stands for ''X is related in a particular way to Y.'' This was needed so that this form of presentation could be used later in the actual experiment. In mixed trial blocks of four trials (for previously trained relations) and then eight trials (for previously derived relations), participants first learned to match a stimulus sequence already acquired (e.g., X1 SAME Y1) to the same stimulus pair presented without a relational cue (e.g., X1–Y1, not X1–Z1). When performance was completely accurate on all previously learned and derived relations in an acquired stimulus set, the same relations were then tested using a new stimulus set in a four-trial (trained relations) and eight-trial (derived relations) block without feedback. When complete accuracy was obtained (12 of 12 trials correct) then they learned with a previously acquired stimulus to match a given stimulus pair (e.g., Z1–X1) to the stimulus sequence that included the relational cue (e.g., pick Z1 OPPOSITE X1, not Z1 SAME X1). Again the participants were tested in the same fashion without feedback with X2, Y2, and Z2 as stimuli, and errorless performance on all trained and derived relations (12 of 12 trials correct) indicated mastery. Failures led to retraining.

Training and Testing Procedures

Four of the successfully pretrained participants went on to the experiment proper. There were three training and testing components in Experiment 1: 1) train four same and opposite relations, and test for analogies between trained and mutually entailed relations; 2) relate one element in each trained pair as same or opposite to a third stimulus and again test for analogies between trained and mutually entailed relations; 3) refresh both forms of training and test for analogies between trained and combinatorially entailed relations.

The stimuli used as sample and comparison stimuli in the training and test phases were two-letter words. There were four sets each consisting of three stimuli: 1) AK (A1), AV (B1), and AF (C1); 2) EW (A2), EJ (B2), and EQ (C2); 3) IH (A3), IZ (B3), and IP (C3); and 4) OC (A4), OX (B4), and OG (C4) (see Figure 1). The feedback procedure and the way of presenting the stimuli were the same as in the computer components of pretraining, with the exception of two new trial types in testing and the inclusion of all four B or C comparison stimuli in the SELECTION RELATA training trials.

Train four A–B relations. Participants learned the following relations: A1 is the same as B1, A2 is the same as B2, A3 is the opposite of B3, and A4 is the opposite of B4. Participants learned each relation one at a time using each relational task (SELECTING RELATA, SELECTING RELATION, PRODUCING RELATA). For example, participants learned to pick B1 (not B2 or B3) in the presence of A1 and SAME (SELECTING RELATA), to pick SAME and not OPPOSITE given A1 ...? B1 (SELECTING RELATION), and produce the correct comparison to A1 SAME ...? (PRODUCING RELATA). In the same way the other three relations were learned. All four relations were then presented mixed in four trial blocks, one for each relational task type, first with feedback, and then without. Complete accuracy led to testing (errors led to retraining).

Find analogies based on trained relations and mutual entailment. Participants were tested for derived analogous relations between both trained relations and derived relations of mutual entailment (note that mutual entailment had never been explicitly tested with these stimuli). Two types of tests were used following this first training phase: selection-based and production-based. Figure 3 presents screen examples for the two new testing tasks. In the selection-based test, participants selected among different pairs of relata as comparisons (SELECTING PAIRS-OF-RELATA). For example, A1–B1 was presented as a sample and A2–B2 or A3–B3 as comparisons (see top of Figure 3). Selecting A2–B2 indicated an analogy based on a trained relation. Similarly, participants were presented with B4–A4 and B2–A2 and B3–A3 as comparisons. Selecting B3–A3 indicated an analogy based on a mutually entailed relation. The test consisted of eight possible trials: four with A–B and four with B–A samples, representing analogies based on trained and mutually entailed relations, respectively. Both types were intermixed, although they are broken out in the results section for theoretical clarity (see Figure 4).

If the participant responded 100% correctly in a trial block (eight trials), a production-based test was given; if not, the selection-based test was given a second time. After the second presentation of the selection test, the production test was given regardless of performance. In the production test, no comparisons were given. Participants produced a pair of relata that bore a similar relation between its members as that of a given sample by typing in the letters of the stimuli (PRODUCING PAIRS-OF-RELATA) in response to a screen

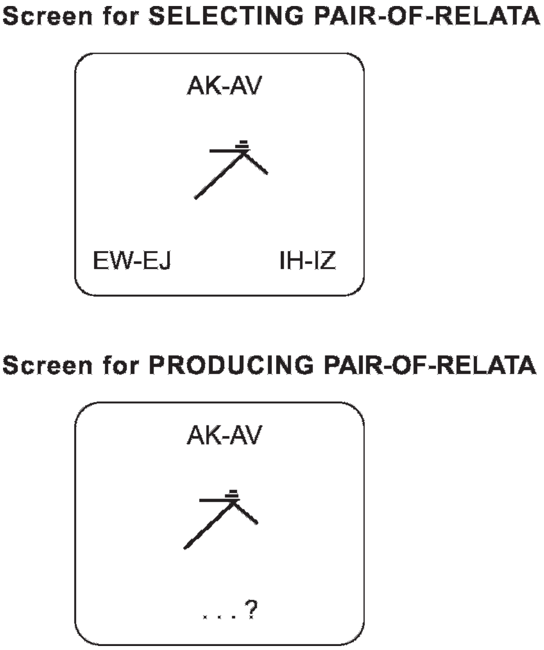

Fig. 3. Examples of the screen arrangements for the two new testing trial types presented in Experiment 1. Stimuli can be linked to sets and relations using Figure 1. In the example of the SELECTING PAIR-OF-RELATA screen, A1–B1 is the sample, and the SAME relational cue sets the occasion for selecting either A2–B2 or A3–B3. In the PRODUCING PAIR-OF-RELATA example, A1–B1 is the sample, and in the presence of the SAME relational cue, a pain of relata must be typed in.

that said, for example (see bottom of Figure 3), A1–B1 SAME ...?. The test consisted of eight possible trials. Four assessed for analogies based on trained relations, with the samples A1–B1, A2–B2, A3–B3, A4–B4, and four for analogies based on mutually derived relations, B1–A1, B2–A2, B3–A3, and B4–A4.

If the participant did not respond correctly to all eight trials in a trial block, the production test was given again. After two presentations of the production test, the A–C training phase started if the participant had made no errors (8/8 in both tests), otherwise the A–B training phase started again. After going back to the A–B training phase a maximum three times, meeting the final criterion or not, the A–C training phase was started.

Train and test A–C relations. In these phases each A stimulus was related as same or opposite to a third (''C'') stimulus, presumably resulting in four 3-stimulus networks. There were three variations on this theme for each participant so that later testing for analogies could combine SAME and OPPOSITE relations in different ways across participants. One participant (Participant 1) picked C stimuli in the presence only of the SAME cue (i.e., A1 SAME C1; A2 SAME C2, and so on); two others (Participants 21 and 4) did so only in the presence of the OPPOSITE cues (i.e., A1 OPPOSITE C1; A2 OPPOSITE C2, and so on); another (Participant 7) received a mixed presentation that combined same and opposite relations with A stimuli from each type of A–B relation previously learned (A1 SAME C1; A2 OPPOSITE C2, A3 SAME C3; A4 OPPOSITE C4). Training and testing were identical to the sequence followed in training and testing A–B relations.

The eight training problems of the A–B and A–C training phase then were presented in mixed order, without feedback, using both kinds of relational tasks (SELECTING PAIRS-OF-RELATA, and PRODUCING PAIRS-OF-RELATA). If the participants performed without error (16 of 16) then the combinatorial test phase started (errors in principle would have led to a repeat of the A–C training phase but this contingency was not contacted).

Find analogies based on combinatorial entailment. The four sets of three A–B and A–C stimuli should presumably have given rise to combinatorially entailed B–C relations. Like the mutually entailed relations previously, these had never been tested prior to analogies testing. The differences in the A–C phase of training resulted in networks composed of one OPPOSITE and one SAME relation (all participants had two networks of this kind), and networks based on two identical SAME/SAME or OPPOSITE/OPPOSITE relations (one participant each had two SAME/SAME or two OPPOSITE/OPPOSITE relations, while a third had one of both). The mixed relation networks yield opposite combinatorial relations while both types of identical relation networks (SAME/SAME or OPPOSITE/OPPOSITE) yield same combinatorial relations. Thus, all participants should have derived two same and two opposite combinatorial relations, but the specific training paths and thus the combination of SAME and OPPOSITE cues that would lead to those relations differed.

The first eight-trial SELECTING PAIRS-OF-RELATA test examined whether analogous

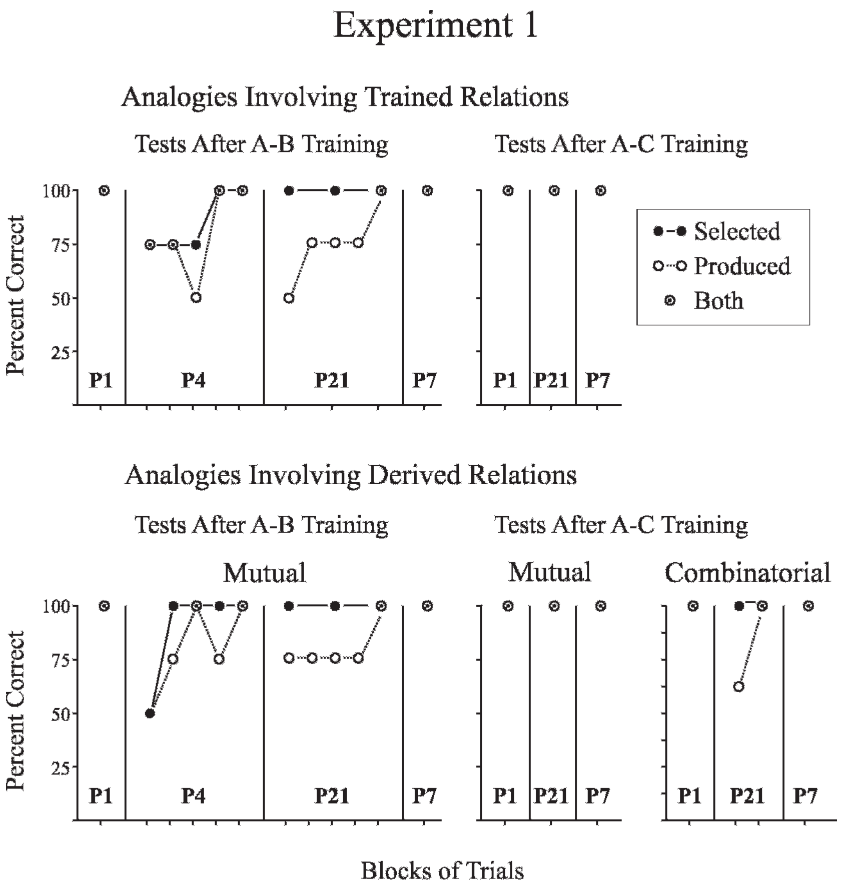

Fig. 4. Derivation of analogies between trained and derived relations in Experiment 1 as shown by the percent ''correct'' in unreinforced testing trials. The size of trial blocks varies (see text); each data point represents four trials except data for analogies through combinatorial entailment which presents eight trials per data point. Note that the upper and lower panels break the same tests into data relevant to analogies through trained and derived relations.

relations were derived based on combinatorially entailed relations. Each participant was tested with each stimulus pair presented in both directions (B1–C1, B2–C2, B3–C3, B4–C4, C1–B1, C2–B2, C3–B3, and C4–B4), using one correct and one incorrect comparison. In the subsequent PRODUCING PAIRS-OF-RELATA test, no comparisons were given, and again each stimulus pair was presented in both directions (a total of eight trials) as a sample, and participants had to type in analogous stimulus pairs. When the participant showed errorless performance in both tests (16 of 16) or when the participant did not meet the final criterion after going back to the mixed training phase three times, the experiment was finished.

Results and Discussion

All 4 participants in Experiment 1 reached the 100% accuracy criterion for all trained relations. The minimum number of training trials presented before the first cycle of the next test phase could start was 64 trials in the A–B training phase, 56 in the A–C phase, and 56 in the mixed training phase. Participant 1 had 74, 56, and 56 trials, Participant 21 had 74, 56, and 58 trials, and Participant 7 had 80, 56, and 58 trials in these phases, respectively. Participant 4 had 114 trials in the A–B training phase after which she withdrew.

In testing, all participants eventually responded 100% correctly on test trials, but the number of retraining cycles before this criterion was reached differed (see Table 1). The more detailed results are shown in Figure 4, with the percent ''correct'' presented for each types of relation within each testing block. Participants 1 and 7 showed analogies based on trained, mutual, and combinatorial relations, never making an error, either in the selection-based or production-based relational testing tasks, in any phase of the experiment. Participant 21 passed all selection-based testing tasks, but needed two reviews of the trained A–B relations to pass the production-based task that assessed for analogies based on trained relations and mutual entailment. When these same performances were assessed in the context of the A–C relations, performance was errorless, as was testing for analogies based on combinatorial entailment. Participant 4 (who, like Participant 21, received all OPPOSITE based A–C relations) failed to derive analogies between the trained relations (A–B) and between the derived relations of mutual entailment (B–A). After two reviews of the trained A–B relations, she responded correctly on all test trials. After this session, she withdrew from the study.

It should be noted that in the PRODUCING PAIRS-OF-RELATA type test trials, more than one correct answer was possible. For example, for Participant 7 the problem B1–C1 SAME ...? presented in the third test phase could be answered correctly by typing any of the follow stimulus pairs: B4–C4, C4–B4, B1–C1, C1–B1, A1–B1, B1–A1, A1–C1, or C1–A1. None of the participants produced correct answers either by repeating the sample or by producing a stimulus combination belonging to the same network (e.g., no one answered A1–C1 to the item B1–C1 SAME ...?). Instead, all productive correct answers were analogies that combined two previously unrelated relational networks on the basis of a common trained, mutual, or combinatorial relation of sameness or opposition.

All participants derived analogies based on trained and mutually entailed relations. All except the participant who withdrew also showed analogies based on combinatorial entailment. The present study is the first in the behavior-analytic literature to show that derived analogous relations can lead to characteristic response topographies. This is important given that productive relational responding has proven difficult for many subjects even with equivalence relations (e.g., Polson et al., 1997; Polson & Parsons, 2000), and the demonstration that analogies based on different topographies are learned and retained differentially (Klein et al., 2007). If one wishes to use derived relational responding as a means of understanding human language and cognition, productive measures are critical since many if not most natural language tasks are not selection-based. Selection-based tasks are less well controlled because comparison sets provide possible controlling variables that are not available in productive tasks. In a typical selection-based task participants may select a correct comparison on the basis of exclusion, for example, if the sample is known to differ from the incorrect comparison. Furthermore, the possibility of chance correct responding is far higher when comparisons are available for selection. Conversely, in principle, virtually an infinite number of responses could occur in the PRODUCING PAIRS-OF-RELATA task.

The present experiment used a comparatively rich set of relations and relational tasks, extending previous work to frames of coordination and opposition, and extending the response alternatives that can be used. In addition to making the present set of results more convincing empirically, that very complexity might help explain the relative success of productive responding as compared to previous, much simpler studies, that had difficulties in that regard (e.g., Polson et al., 1997; Polson & Parsons, 2000). From an RFT perspective the key issue is whether the training and testing tasks include adequate relational contextual cues (C_{rel}) since it is argued that mere conditional discriminations should not lead to derived relational responding without cues that evoke relational operants. Based on that reasoning, task simplicity could make it hard to demonstrate productive responding, and a long set of conditional discrimination trials without sufficient relational cues (e.g., Polson et al., 1997) could actually weaken productive relational responding.

Considered in isolation, explanations other than analogy are possible for some of the specific performances seen here, but the A–C training was varied to control for sources of control other than the derivation of relations among relations. For example, consider Participant 1 who had learned A–C relations in the presence of SAME. Given B1–C1 SAME, and the comparisons B2–C2 and B3–C3, selecting B2–C2 might be possible on the grounds that B1 as well as B2 has previously been present when the SAME stimulus was present. Participant 7, however, controls for this possibility. Solving B4–C4 SAME ? with B1–C1 and B3–C3 as comparisons can no longer be based on comparison of the B (or the C) stimuli of the sample with the B (or the C) stimuli of the comparisons. To explain the consistent correct responding in all these trials of the three test phases, an analysis in terms of relating relations seems to be demanded.

EXPERIMENT 2. UNDERSTANDING EXPOSITORY ANALOGIES WITH SAME/OPPOSITE RELATIONS

In an expository analogy, the participants have to find what kind of relation holds among the relata given their relation to relata in another relational network. A common sense example of an expository analogy might be something like ''an atom is like the solar system'' and then questioning how parts of the atom relate. This is similar to the method used in the Relational Evaluation Procedure (Hayes & Barnes, 1997) which has been used in RFT work on analogies that combine same and different relations (Stewart et al., 2004).

Method

Participants, Pretraining, Setting, and Stimuli

The 7 participants were those remaining from pretraining in Experiment 1. The pretraining, stimuli, and setting are identical to that experiment.

Training and Test Procedures

Training A–B relations by analogy. The primary difference between Experiment 1 and Experiment 2 is that participants were trained in part by analogy. In the A–B training, SELECTING RELATA, PRODUCING RELATA, and SELECTING RELATION tasks were used to train A1 SAME B1, A3 OPPOSITE B3. This is identical to Experiment 1. The SELECTING PAIRS-OF-RELATA task was then used to train A2–B2 SAME A1–B1, and A4–B4 SAME A3–B3. The training blocks and criteria were otherwise identical to Experiment 1.

Selecting the relation acquired by analogy. In the test, participants had to select the relational cue that indicated the kind of relation that held between a given pair of relata (SELECTING RELATION). This task applied only to those two relations established by analogy (e.g., given A4 …? B4 choose OPPOSITE not SAME), and to the two mutually entailed relations (e.g., given B2 …? A2 choose SAME not OPPOSITE).

After errorless performance (4/4) the A–C training started. The test was given a second time if errors were made. If after two presentations of the test the criterion was not met, A–B training was given again. The A–C training started when the participant met the errorless test criterion or did not do so after going back to the A–B training three times.

Creating relational networks via analogy. As in Experiment 1, three different kinds of networks were established by varying the specific A–C relations trained. In this experiment, the focus was on creating networks that were either all SAME or all OPPOSITE or mixed, but based on analogy as a method of acquisition. In a fashion identical to the first phase of training, Participants 2 and 22 learned A1 SAME C1, A3 SAME C3, A2–C2 SAME A1–B1, and A4–C4 SAME A1–B1; Participants 5 and 10 learned A1 OPPOSITE C1, A3 OPPOSITE C3, A2–C2 SAME A3–B3, and A4–C4 SAME A3–B3; and Participants 8, 18, and 20 learned A1 SAME C1, A3 SAME C3, A2–C2 SAME A3–B3, and A4–C4 SAME A3–B3. Training blocks and procedures were otherwise identical to the parallel training in Experiment 1. The test was identical to the previous test, except that it focused on A–C relations instead of A–B relations.

All the A–B and A–C training problems were then presented in mixed order in eight trial blocks, without feedback. If the participants performed without error (8/8) then the combinatorial (B–C and C–B) test started (errors led to a repeat of the A–C and mixed training, up to three times). In the final testing, all four combinatorially entailed relations that could have been acquired by analogy were tested in four trial blocks. The test was given a second time if any errors were made. If after two presentations the criterion was not met, the mixed training phase was given again. The experiment was finished when the participant met the final test criterion or when the participant did not meet the final criterion three times.

Results and Discussion

All 7 participants in Experiment 2 reached the 100% accuracy criterion for all trained relations. The minimum number of training trials presented before the test phase could start was 56, 48, and 48 trials respectively for the A–B, A–C, and mixed training phases. Participant 2 had 74, 64, and 170 trials; Participant 22 had 102, 132, and 72 trials; Participant 5 had 70, 90, and 142 trials; Participant 10 had 62, 56, and 54 trials; Participant 8 had 82, 52, and 56 trials; Participant 18 had 62, 52, and 92 trials; and Participant 20 had 98, 58, and 54 trials respectively in these phases.

Table 1 shows the results expressed in terms of testing to criteria; Figure 5 shows the more detailed percentages of ''correct'' responding across testing trial blocks. In testing, Participant 22 was errorless throughout. Participants 10 and 20 passed all tests as well after one or two (respectively) exposures to the A–B training phase, and two exposures to the mixed A–B A–C training phase. Participant 2 failed the A–B training and testing phase but passed the identical A–C testing phase, and the combinatorial testing phase without error. The remaining 3 participants (5, 8, and 18) all met the criterion in the test phase based on A–B training, but failed to do so in the identical phase based on A–C training. After three failures they were passed on to the mixed training phase anyway but after three cycles did not achieve 100% accuracy.

The majority of participants passed each phase of the experiment, but the results are considerably weaker than in Experiment 1. We were concerned that the repetitive nature of the task could be part of the problem. There was only one kind of test trial in Experiment 2, SELECTING RELATION, and there were only two comparison stimuli, the SAME and the OPPOSITE stimulus, to choose from. Of the 7 participants, 6 met criterion in the first (A–B) phase but only 4 did so in the structurally identical (A–C) second phase, suggesting that participants were becoming less engaged in the task. One response could have been to simplify the task further, but as with our interpretation of the previous failures of productive responding, excessive simplicity could actually undermine the relational nature of the task. Increasing motivation for correct performances would have been another alternative (especially given that participants were college students working for credit) but that would have introduced an additional variable. Instead, we simplified the stimulus sets but crafted a more challenging test in the next two experiments that was similar to Experiments 1 and 2 but was based on nonsymmetrical comparative relations.

Table 1

The number of training cycles (1, 2, or 3) needed to meet the criterion for each participant in each test in each experiment.

		Experiment 1 Tests		
		A-B	A-C	Combinatorial
#1	S:	1	1	1
	P:	1	1	1
#4	S:	2	Participant withdrew	
	P:	3		
#21	S:	1	1	1
	P:	3	1	1
#7	S:	1	1	1
	P:	1	1	1
		Experiment 2 Tests		
		A-B	A-C	Combinatorial
#2	R:	-	1	1
#22	R:	1	1	1
#5	R:	2	-	-
#10	R:	2	1	3
#8	R:	2	-	-
#18	R:	1	-	-
#20	R:	1	1	3
		Experiment 3 Tests		
#12	S:	1		
	P:	1		
#28	S:	-		
	P:	-		
#30	S:	3		
	P:	3		
		Experiment 4 Tests		
#19	R:	2		
#34	R:	1		

-: Participant failed to meet the criterion
S: SELECTING PAIRS-OF-RELATA
P: PRODUCING PAIRS-OF-RELATA
R: SELECTING RELATION

EXPERIMENT 3: FINDING ANALOGIES WITH SAME/SMALLER/ LARGER RELATIONS

In Experiment 1 two kinds of relations were used, same and opposite, both of which are

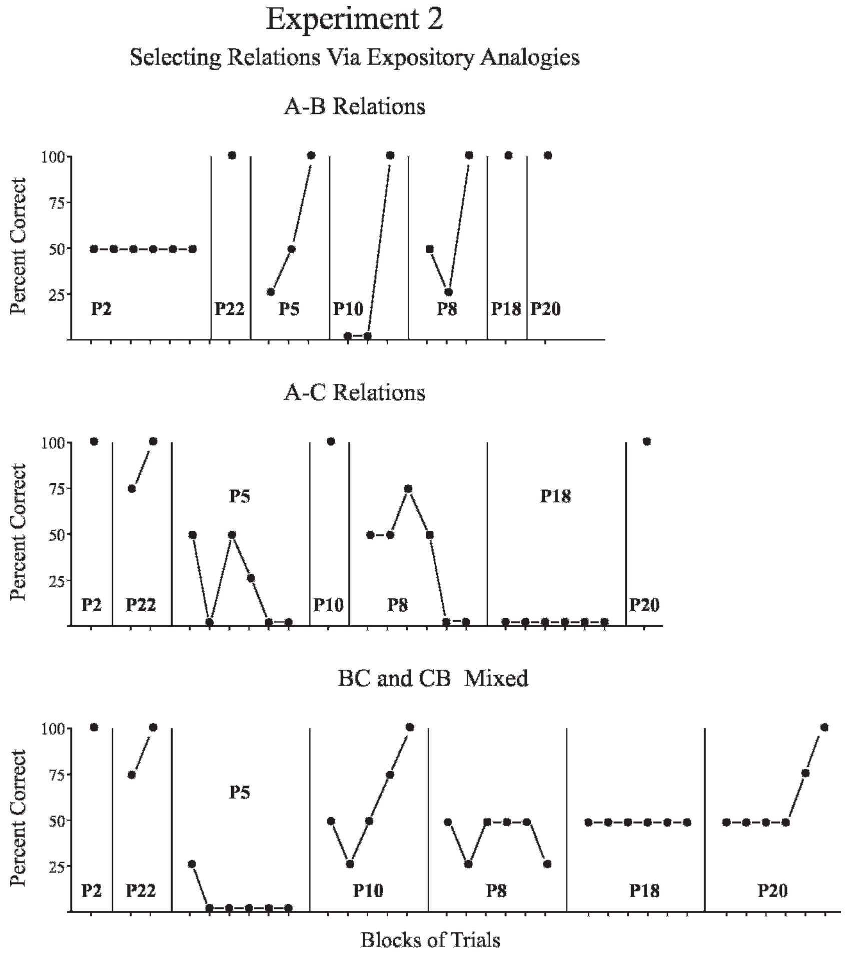

Fig. 5. Relations derived by expository analogies in Experiment 2. The size of trial blocks vary (see text).

symmetrical. In Experiment 3 the relations of same, "is smaller than", and "is larger than" were used to see whether analogies can be found based on nonsymmetrical relations using both selection-based and topography-based tests. None of these performances have yet been shown in the behavior-analytic literature.

In Experiment 1 participants were taught four different relational networks and then asked to select or produce a pair of relata from one network that bore a similar relation to another pair of relata in one of the other three networks. Similarly, in Experiment 3 participants chose which part of another network was relevant and which was not for the production of an analogy.

Method

Participants, Pretraining, Setting, and Stimuli

Five participants completed pretraining; 3 were assigned to Experiment 3 and 2 to Experiment 4. The setting, apparatus, materials and instructions were identical to Experiment 1 except that in pretraining OPPOSITE was not trained. Pretraining was identical to Experiment 1 except that SMALLER THAN and LARGER THAN relational cues (see Figure 1) were trained using the three comparison stimulus sets. For example, when the small, medium, and large circle was used as a comparison set, the medium circle was used as a sample and selecting the smaller circle in the presence of the SMALLER THAN cue, selecting the larger in the presence of the LARGER

THAN cue, and selecting the medium circle in the presence of SAME was reinforced.

Training and Testing Procedures

The stimuli used as sample and comparison stimuli in the training and test phases were two-letter words. There were two sets, each consisting of three stimuli: 1) AK (A1), AV (B1), and AF (C1); and 2) EW (A2), EJ (B2), and EQ (C2). In a fashion identical to the A–B training in Experiment 1, participants learned a smaller set of A–B and A–C relations. Participant 12 learned A1 SAME B1, A1 SMALLER THAN C1, A2 SAME B2, and A2 SMALLER THAN C2 using SELECTING RELATA, SELECTING RELATION, and PRODUCING RELATA tasks. Participants 28 and 30 learned a more complex pair of networks: A1 SAME B1, A1 SMALLER THAN C1, A2 SAME B2, and A2 LARGER THAN C2.

Participants were tested for derived analogous relations between trained relations, derived relations of mutual entailment, and combinatorial entailment. The participants had to choose which relation within a second network was relevant for the production of an analogy with a relation of the first network using the SELECTING PAIRS-OF-RELATA and PRODUCING PAIRS-OF-RELATA tasks as in Experiment 1.

Participant 12 (with networks both based on SAME and SMALLER THAN) was tested for the following six analogous relations (tested in both directions, for 12 total trials): A1–B1 and A2–B2, A1–C1 and A2–C2, B1–A1 and B2–A2, C1–A1 and C2–A2, B1–C1 and B2–C2, and C1–B1 and C2–B2. Participants 28 and 30 (with networks based on both SMALLER THAN and LARGER THAN stimuli) were tested for the following six analogous relations in both directions, also for 12 total trials): A1–B1 and A2–B2, A1–C1 and C2–A2, B1–A1 and B2–A2, C1–A1 and A2–C2, B1–C1 and C2–B2, and C1–B1 and B2–C2. For all participants, the incorrect comparison for trained, mutual, or combinatorial relations was the other available trained, mutual, or combinatorial relation. The samples for the PRODUCING PAIRS-OF-RELATA trails consisted of all possible stimulus pairs in both networks.

Note that the difference between these two patterns of training resulted in more complex derived relations for Participants 28 and 30. Participant 12 was asked to derive analogies between trained relations (e.g., between A1–B1 and A2–B2), between derived relations of mutual entailment (e.g., between B1–A1 and B2–A2), and between derived relations of combinatorial entailment (e.g., between B1–C1 and B2–C2). Participants 28 and 30 had to find analogies between trained relations and derived relations of mutual entailment (e.g., between B1–A1 and A2–B2), between trained relations and derived relations of combinatorial entailment (i.e., between C1–B1 and A2–C2), and between derived relations of mutual entailment and combinatorial entailment (i.e., between B2–C2 and C1–A1).

Each type of test (selecting and producing) consisted of 12 trials. The test was given a second time if one or more errors were made. If after a maximum two presentations of each test the criterion was not met (12/12 in both tests) the training was given again. The experiment ended when the test criterion was met or when the participant did not meet the final criterion after going back to the first training three times.

Results and Discussion

All 3 participants in Experiment 3 reached the 100% accuracy criterion for all trained relations. The minimum number of training trials presented before the first test was 64 trials. Participant 12 had 76, Participant 28 had 128, and Participant 30 had 100 trials before reaching the criterion.

The results are shown in Table 1 and Figure 5. Participant 12 (who received SAME and SMALLER THAN training) responded errorlessly on all test trials. Participant 30 achieved 100% accuracy in the third training cycle. After three reviews of the training, Participant 28 failed to derive analogies.

As in Experiment 1, in the PRODUCING PAIRS-OF-RELATA test, more than one correct answer was possible. For example, for Participant 12 the problem B1–C1 SAME ...? presented in the third test could be answered correctly by typing B2–C2, A2–C2, A1–C1, or B1–C1. Although just repeating the sample would always result in a correct score, none of the participants ever did so, nor did they ever produce a stimulus combination belonging to the same network (e.g., B1–C1 SAME ...? A1–C1).

Although 1 of the 3 participants did not meet the criterion, the performances of the other participants are so elaborate (since they required 24 of 24 trials correct including a productive and not just a selective response) that they seem unexplainable without appealing to relations among relations as the basis of responding.

EXPERIMENT 4: UNDERSTANDING EXPOSITORY ANALOGIES WITH SAME/SMALLER/LARGER RELATIONS

In analogy with Experiment 2, participants in Experiment 4 had to find what kind of relation held among relata given their relation to relata in another relational network. Instead of same and opposite, the relations of sameness, ''is smaller than'', and ''is larger than'' were used and two relational networks were trained instead of four. The expository analogies using nonsymmetrical relations have not previously been shown in the behavior-analytic literature.

Method

Participants, Pretraining, Setting, and Stimuli

The remaining 2 participants described in the pretraining section of Experiment 3 served in this study. The setting, apparatus, materials, stimuli, and pretraining were identical to Experiment 3.

Training and Testing Procedures

In a fashion identical to Experiment 2, Participant 19 learned A1 SAME B1, A1 SMALLER THAN C1, A2–B2 SAME A1–B1, and A2–C2 SAME A1–C1. Participant 34 learned the same set except the last relation was reversed: A2–C2 SAME C1–A1. As in Experiment 2, in the test participants had to select what kind of relation held between a given pair of relata: same, smaller than, or larger than (SELECTING RELATION). The following relations were tested: A2 ...? B2, A2 ...? C2, B2 ...? A2, C2 ...? A2, B2 ...? C2, and C2 ...? B2. The test was given a second time if one or more errors were made. If after two presentations the criterion was not met the training was given again. The experiment was finished when the participant met the final test criterion (6/6) or when the participant did not meet the final criterion after going back to the training three times.

Results and Discussion

Both participants reached a 100% accuracy criterion for all trained relations. The minimum number of training trials presented before the first cycle of the test phase could start was 56 trials. Participant 19 had 70 and Participant 34 had 68 trials. Results are shown in Table 1 and Figure 6. Participant 19 met the test criterion after two training cycles. Participant 34 responded without error to all test trials. Since there were always three comparisons present, SAME, SMALLER THAN, and LARGER THAN, it seems likely that correct responding was controlled by the expository analogy— chance alone would have provided little opportunity for consistent success.

GENERAL DISCUSSION

The present study showed that participants were able to find an analogy between a relation in one set and a relation in another set and when exposed to an analogy were able to derive particular relations among new events. This process occurred both with symmetrical relations (sameness, opposition) and with nonsymmetrical relations (smaller than, larger than), and was demonstrable with tasks that required participants to select relata, select related pairs, select relational cues, and perhaps most importantly, with tasks that required participants to produce either individual relata or related pairs.

Not all participants showed analogous relational responding. Results were weakest in Experiment 2, which had a simple two-choice response task, particularly as the experiment progressed beyond the initial stages. But each of the key features of an RFT approach to analogy received support when all participants in all experiments are considered. Especially convincing are the productive performances, since accidental sources of control by the comparisons are absent and a virtually infinite variety of response forms are possible.

This study was meant to see if experimental procedures occasioned the emergence of certain kinds of relational responses that are consistent with an RFT account of analogical responding, going beyond what has been previously demonstrated in a behavior-analytic laboratory setting. This is not the same as

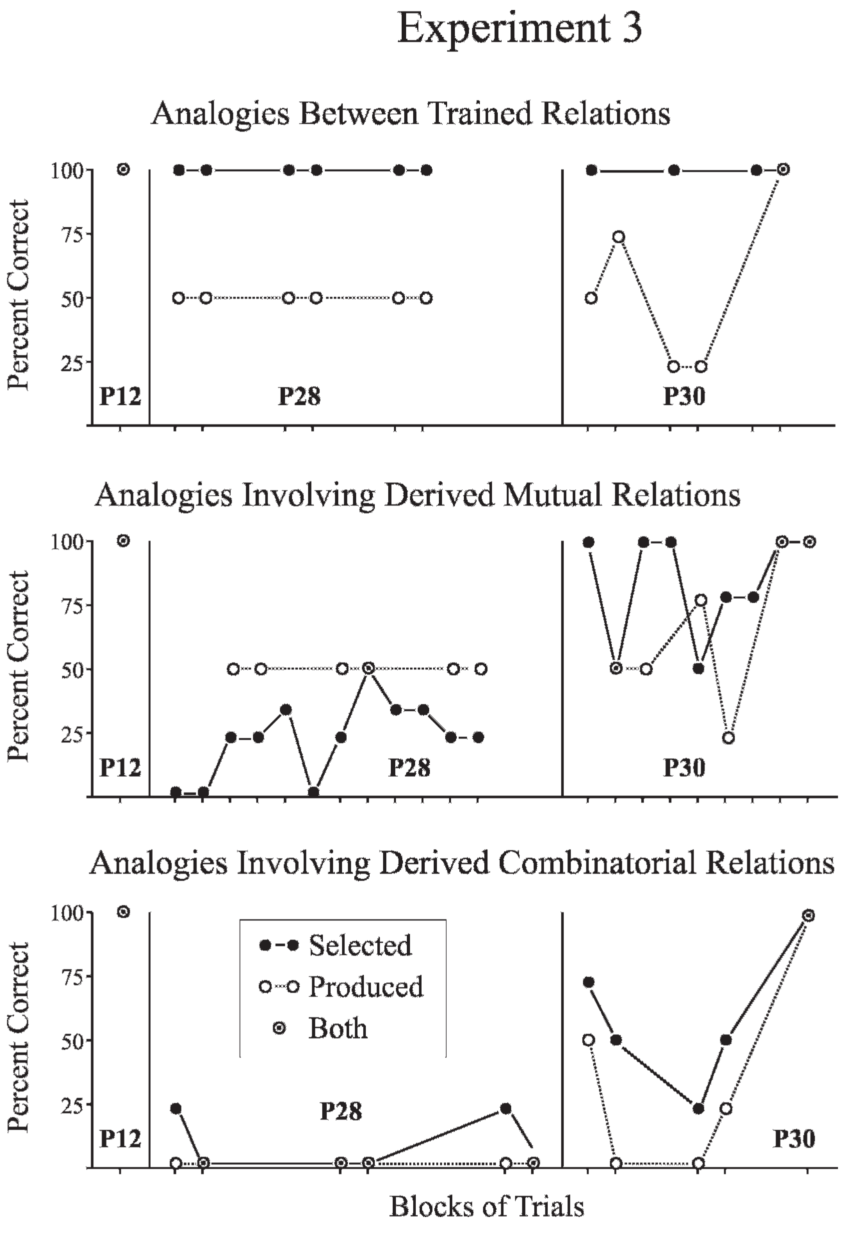

Fig. 6. Derivation of analogies between trained and derived relations in Experiment 3. The size of trial blocks vary (see text).

experimentally analyzing all of the sources of derived relational responding that are relevant to the present study, since the performances here depend upon relational abilities that go far beyond the present study. The task of the experimental analysis of these component skills *has* been undertaken in the overall RFT research program, however. Considerable progress has been made demonstrating that derived relational responding is an operant initially based on shaped multiple exemplars (e.g., Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Barnes-Holmes, Barnes-Holmes, Smeets, Strand, & Friman, 2004; Berens & Hayes, 2007; Luciano et al., 2007). For example, Berens and Hayes (2007) showed

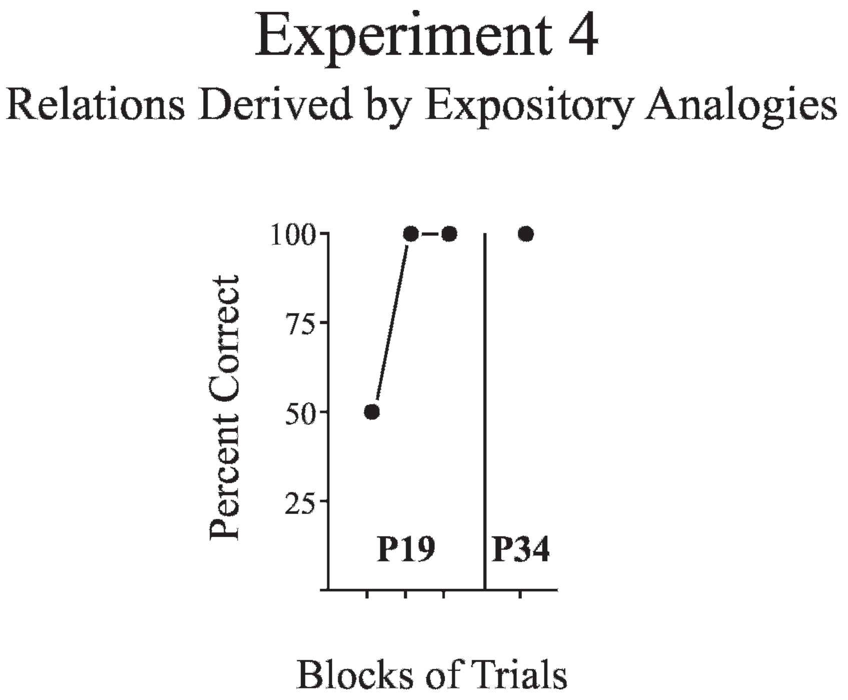

Fig. 7. Relations derived by expository analogies in Experiment 4.

that a history of reinforcement with both nonarbitrary comparative relations and arbitrary comparative relations was responsible for the development of arbitrarily applicable derived comparative relational responding. Luciano et al. (2007) provided similar data for frames of coordination, or equivalence. When these data are combined with the present study and other existing behavioral literature on analogy (e.g., Barnes-Holmes et al., 2005), the outlines of a relatively complete experimental analysis of analogy are now apparent.

Ambiguity in the conditions under which an experimental analysis can be said to have been accomplished raises a particular difficulty regarding scaling behavior analytic research strategies that were initially developed with nonhuman animals into complex human verbal and cognitive performances. It is relatively common to see criticisms (e.g., McIlvane, 2003) that specific RFT research studies often fail to deal with preexperimental histories and thus the complex behavior seen in RFT experiments could be due to unknown properties of human language. Taken to its extreme, this would mean that experimentally adequate basic behavior-analytic research on complex human behavior is impossible, since in dealing with performances such as analogy, problem-solving, or reasoning it is not practical to combine an experimental analysis of all of components of complex performances into a single study. Frames of coordination appear to be acquired as an operant in infancy (Luciano et al., 2007). Comparative relational operants appear to be acquired in the 3- to 5-year-old range (Berens & Hayes, 2007). Analogies build on such performances. Given the age range and the necessarily uncontrolled nature of early childhood education and experiences, only fantasy experiments, not one likely to occur in the real world, would be able to conduct adequately controlled experimental analyses of all of these behaviors at once. The RFT solution is to conduct inductive experimental analyses of each element of a complex account within the entire RFT research program. So far as we know, no other strategic approach has been proposed for how to conduct an experimental analysis of elaborate verbal and cognitive performances. Perhaps for this reason, interpretation is often left as the seemingly ‘‘conservative’’ alternative (e.g., Palmer, 2004)—but this is an odd place to end up in a tradition that calls itself the ''experimental analysis of behavior.''

Another specific example of the same problem is the fact that 6 participants were eliminated from the study because they failed to acquire control by features of the task that would be required to assess the actual questions of interest (e.g., acquire relational cues, demonstrate their arbitrary applicability). Because these are processes that are by now very well established in the literature and themselves have been subject to experimental analysis, setting aside these participants on practical grounds seem justified given the focus of the research. It is another example, however, of how traditional behavior analytic research strategies may need to be adjusted in order to meet the practical realities of experimental research on complex verbal and cognitive performances.

Because analogies are traditionally based on frames of coordination between sets of relations, it might be argued that equivalence might alone account for the present results, provided only that relations are admitted as relata. This possibility does not seem to withstand close scrutiny because it does not explain how derived relations could in fact be treated as relata in that way without an appeal to multiple relational operants or their equivalent. Consider two trained networks: A1 SMALLER THAN B1, A1 SAME C1 and A2 GREATER THAN B2, A2 SMALLER THAN C2. As is shown by the present experiments, given that training history, a normal adult participant would be likely to derive an equivalence relation between B1–C1and C2–B2, since it can be derived that B1 is greater than C1 and C2 is greater than B2. The source of that performance could not be the relational cues alone, however, which differ completely in the two training sets. It could not be the relata alone since one must reverse the B–C relation in one network as compared to the other to create the analogy. Thus, one would need to explain how the relata in combination with SAME, SMALLER THAN, and GREATER THAN came to exert control over the equivalence relation between B1–C1 and C2–B2. As we have pointed out previously, it is not progressive to avoid the concept of relational frames merely by speaking vaguely about stimulus control or stimulus generalization but without providing the history needed for relational stimulus control as it exists to emerge (Hayes & Barnes, 1997). Thus far no one within behavior analysis has yet provided an alternative and experimentally viable solution to this problem, essentially leaving the field to RFT researchers in the empirical analysis of such phenomena. Any adequate account we have been able to imagine for relational stimulus control simply replicates the RFT account in different terms.

Empirical analysis, not mere interpretation, seems needed if behavior analysis is to enter an era of the experimental analysis of language and cognition. Analogy is a good place to begin since ''Differential psychologies have long recognized the close relation between analogical reasoning and intelligence'' (Sternberg, 1977b, p. 353). Derived relational responding in general seems to be a key component of intellectual behavior (e.g., O'Hora, Pelaez, & Barnes-Holmes, 2005), but from an RFT point of view analogy should be particularly so, because relating sets of relations allows for the rapid development of elaborated and subtle forms of verbally established stimulus control. For example, the statement ''anxiety is like quicksand'' may rapidly bring an entire set of responses relevant to the situation (what to do when confronted with an aversive situation in which escape responses produce further harm). Without analogy and metaphor relational networks would need to be built up, piece by piece, in every stimulus situation of relevance. This may be one reason most abstract concepts are themselves originally analogies and metaphors (Skinner, 1989).

There is much more to do in the basic RFT analysis of analogy and metaphor, such as seeing if these processes produce transformations of stimulus functions, and understanding how they are regulated by nonarbitrary features of the environment in pragmatic verbal analysis. Given the known importance of analogical reasoning, however, applied research on the establishment of good analogical skills may be of special importance. Cognitive psychologists have dominated in the research literature in the area of analogy and metaphor, but their work has led to only modest advances in the promotion of analogical reasoning. In this area, cognitive psychology has relied on vague concepts such as ''knowledge'' in ''memory'' that is ''transferred'' from one ''domain'' to another (e.g., Vosniadou & Ortony, 1989), and that are explained through an appeal to hypothesized internal mechanisms, such as ''analogy machines'' that ''map'' (e.g., Holyoak & Thagard, 2002; Keane, Ledgeway, & Duff, 1994), and their neurobiological substrates. In contrast, RFT approaches the area as a matter of contextually controlled learned relational responses, in which networks of stimulus relations are related one to the other based on history and context. Interestingly, cognitive scientists themselves are beginning to model exactly these kinds of performances and to argue their centrality in human language and cognition, but so far these more experiential and contextual models are dominantly based on computer simulations, not actually experimental histories (e.g., Doumas et al., 2008). Without that step, applied programs seem impossible.

Good behavioral technologies exist for teaching concepts based on nonarbitrary examples (e.g., Engelmann, Becker, Carnine, & Gersten, 1988), and progress is being made in applying behavioral technology to the instruction of arbitrarily applicable relational responses (Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Berens & Hayes, 2007; Rehfeldt & Barnes-Holmes, in press). Just as in other areas of RFT research that seem to be shifting to the applied area (Rehfeldt & Root, 2005), the present study suggests that applied behavior analysts might approach the training of this important skill from an RFT point of view by focusing on how to develop the arbitrarily applicable relational core of this performance. In the real world arbitrarily applicable relational responses are rarely arbitrarily applied. Consider the question: ''which is larger, the earth or the sun?'' The formal properties of the related events determine the answer (''the sun''), but only based on actions that could as well be applied to ''a > b > c'': a simple nonarbitrary comparative relation (directly examining the earth and sun and deriving a comparative relation) will not lead to the correct answer. Thus, the comparative relations involved are arbitrarily applicable, but in this instance they are not arbitrarily applied. If this analysis is correct, then much more emphasis should be placed on the arbitrary applicability of relational responding underlying analogy. Even preschool children can respond to games that demonstrate the arbitrary nature of relational activity (''if more was less and less was more, which would you want, more candy or less?'' or ''If I were your son, who would you be?''). Perhaps early training programs that emphasized the abstract nature of relational responding would be useful in developing the relational skills that RFT argues underlie analogical reasoning.

Mainstream cognitive science might also benefit methodologically from key features of a RFT approach, even for testing traditional cognitive theories. Cognitive studies of analogy have almost all relied on natural language examples because the core processes underlying language and cognition are themselves not well specified. RFT specifies the nature of analogous relations and thus can be used to construct highly specific empirical models of complex cognitive phenomena that require no appeal to natural language. This study addresses analogical reasoning in an entirely abstract context in which the history with regard to the specific relata and the specific relations among them is relatively well controlled. There is no reason to rely solely on analogies and metaphors drawn from day to day language as the only experimental preparation—with all the attendant problems such an approach brings in terms of cultural bias, the possibility of previous nonrelevant histories, and the inability to compare performances across species. The present method could be used virtually without modification with college students, elementary school children, or chimpanzees, whether in the United States or Mozambique, and thus may provide a more uniform and less biased preparation in which to look at the development and impact of analogical performances and their neurobiological correlates.

Behavioral psychology has not been at the forefront in the experimental analysis of complex human behavior such as reasoning and thinking. The present study provides a small additional indication that this barrier is more a matter of needed research effort than the absence of adequate and applicable behavioral concepts and methods.

REFERENCES

Aristotle. (1941a). *Ethica Nicomachea* (W. D. Ross, trans.).In R. McKeon (Ed.), *The basic works of Aristotle* (pp. 927–1112). New York: Random House.

Aristotle. (1941b). *Metaphysics* (W. D. Ross, trans.).In R. McKeon (Ed.), *The basic works of Aristotle* (pp. 680–926). New York: Random House.

Aristotle. (1963). *Metaphysics* (A. E. Wardman & L. L. Creed, trans.)In R. Bambrough (Ed.), *The philosophy of Aristotle* (pp. 31–132). New York: New American Library.

Barnes, D., Hegarty, N., & Smeets, P. M. (1997). Relating equivalence relations to equivalence relations: A relational framing model of complex human functioning. *The Analysis of Verbal Behavior, 14*, 57–83.

Barnes-Holmes, Y., Barnes-Holmes, D., & Smeets, P. M. (2004). Establishing relational responding in accordance with opposite as generalized operant behavior in young children. *International Journal of Psychology and Psychological Therapy, 4*, 559–586.

Barnes-Holmes, Y., Barnes-Holmes, D., Smeets, P. M., Strand, P., & Friman, P. (2004). Establishing relational responding in accordance with more-than and less-than as generalized operant behavior in young children. *International Journal of Psychology and Psychological Therapy, 4*, 531–558.

Barnes-Holmes, D., Regan, D., Barnes-Holmes, Y., Commins, S., Walsh, D., Stewart, I., et al. (2005). Relating derived relations as a model of analogical reasoning: Reaction times and event-related potentials. *Journal of the Experimental Analysis of Behavior, 84*, 435–451.

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.

Brown, A. L. (1989). Analogical learning and transfer: What develops? In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 369–412). Cambridge: Cambridge University Press.

Clement, J. (1988). Observed methods for generating analogies in scientific problem-solving. *Cognitive Science, 12(4)*, 563–586.

Doumas, L. A. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review, 115*, 1–43.

Edwards, P. (Ed.) (1967). *Encyclopedia of Philosophy* (Vols. 1–8). New York: The Free Press.

Engelmann, S., Becker, W. C., Carnine, D. W., & Gersten, R. (1988). The direct instruction follow through model: Design and outcomes. *Education and Treatment of Children, 11*, 303–317.

Euclid. (2d edition, 1956). Book V (trans. from the text of Heiberg). In T. L. Heath (Ed.), *The thirteen books of Euclid's elements* (Vol. II, p. 112–186). Cambridge: Cambridge University Press.

Eysenck, M. W., & Keane, M. T. (1995). *Cognitive psychology: A student's handbook* (3rd ed.). Hillsdale, NJ: Erlbaum.

Getner, D., & Holyoak, K. J. (1997). Reasoning and learning by analogy: Introduction. *American Psychologist, 52*, 32–34.

Goetz, P. W. (Ed.) (1987). *New Encyclopædia Britannica, Micropædia, Ready Reference.* (15th edition, Vols. 1–32). Chicago: Encyclopædia Britannica.

Hayes, S. C., & Barnes, D. (1997). Analyzing derived stimulus relations requires more than the concept of stimulus class. *Journal of the Experimental Analysis of Behavior, 68*, 235–270.

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

Hayes, S. C., Gifford, E. V., Townsend, R. C., Jr., & Barnes-Holmes, D. (2001). Thinking, problem-solving, and pragmatic verbal analysis. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational Frame Theory: A post-Skinnerian account of human language and cognition* (pp. 87–101). New York: Plenum Press.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought.* Cambridge, MA: MIT Press.

Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist, 52*, 35–44.

Holyoak, K. J., & Thagard, P. (2002). Analogical mapping by constraint satisfaction. In T. A. Polk, & C. M. Seifert (Eds.), *Cognitive modeling* (pp. 849–909). Cambridge, MA: MIT Press.

Johnson, C., & Sidman, M. (1993). Conditional discrimination and equivalence relations: Control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 59*, 333–347.

Kamouri, A. L., Kamouri, J., & Smith, K. H. (1986). Training by exploration: Facilitating the transfer of procedural knowledge through analogical reasoning. *International Journal of Man-Machine Studies, 24*, 171–192.

Keane, M. T., Ledgeway, T., & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science, 18*, 387–438.

Klein, P. D., Piacente-Cimini, S., & Williams, L. A. (2007). The role of writing in learning from analogies. *Learning and Instruction, 17*, 595–611.

Kolodner, J. L. (1997). Educational implications of analogy: A view from case-based reasoning. *American Psychologist, 52*, 57–66.

Lipkens, R. (1992). *A behavior analysis of analogical reasoning.* Paper presented at the meeting of the International Association for Behavior Analysis, San Francisco, California.

Lipkens, R., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of the development of derived relations in an infant. *Journal of Experimental Child Psychology, 56*, 201–239.

Luciano, C., Becerra, I. G., & Valverde, M. R. (2007). The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. *Journal of the Experimental Analysis of Behavior, 87*, 349–365.

McIlvane, W. J. (2003). A stimulus in need of a response: A review of Relational Frame Theory: A Post-Skinnerian account of human language and cognition by Hayes, Barnes-Holmes, and Roche. *The Analysis of Verbal Behavior, 19*, 29–38.

Nelson, D. L., McKinney, V. M., Gee, N. R., & Janczura, G. A. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. *Psychological Review, 105(2)*, 299–324.

Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem-solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17(3)*, 398–415.

O'Hora, D., Pelaez, M., & Barnes-Holmes, D. (2005). Derived relational responding and performance on verbal subtests of the WAIS-III. *Psychological Record, 55(1)*, 155–175.

Ortony, A. (1993). *Metaphor and thought.* Cambridge, England: Cambridge University Press.

Palmer, D. C. (2004). Data in search of a principle: A review of S. C. Hayes, D. Barnes-Holmes, and B. Roche (Eds.), *Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition. Journal of the Experimental Analysis of Behavior, 81*, 189–204.

Polson, D. A. D., & Parsons, J. A. (2000). Selection-based versus topography-based responding: An important distinction for stimulus equivalence? *The Analysis of Verbal Behavior, 17*, 105–128.

Polson, D. A. D., Grabavac, D. M., & Parsons, J. A. (1997). Intraverbal stimulus-response reversibility: Fluency, familiarity effects, and implications for stimulus equivalence. *The Analysis of Verbal Behavior, 14*, 19–40.

Polya, G. (1954). *Mathematics and plausible reasoning: Volume 1. Induction and analogy in mathematics*. Princeton, NJ: Princeton University Press.

Rehfeldt, R., & Barnes-Holmes, Y. (Eds.) (in press). *Applied programs for derived relational responding*. Oakland, CA: New Harbinger.

Rehfeldt, R. A., & Root, S. L. (2005). Establishing derived requesting skills in adults with severe developmental disabilities. *Journal of Applied Behavior Analysis, 38*, 101–105.

Sidman, M. (2008). Symmetry and equivalence relations in behavior. *Cognitive Studies, 15*, 322–332.

Skinner. (1957). *Verbal behavior*. East Norwalk, CT: Appleton-Century-Crofts.

Skinner. (1989). The origins of cognitive thought. *American Psychologist, 44*, 13–18.

Spearman, C. (1923, reprinted 1973). *The nature of 'intelligence' and the principles of cognition*. New York: Arno Press.

Sternberg, R. J. (1977a). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Mahwah, NJ: Erlbaum.

Sternberg, R. J. (1977b). Component processes in analogical reasoning. *Psychological Review, 84*, 353–378.

Stewart, I., Barnes-Holmes, D., Hayes, S. C., & Lipkens, R. (2001). Relations among relations: Analogies, metaphors, and stories. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational Frame Theory: A post-Skinnerian account of human language and cognition* (pp. 73–86). New York: Plenum Press.

Stewart, I., Barnes-Holmes, D., & Roche, B. (2004). A functional-analytic model of analogy using the relational evaluation procedure. *Psychological Record, 54*, 531–552.

Stewart, I., Barnes-Holmes, D., Roche, B., & Smeets, P. (2001). Generating derived relational networks via the abstraction of common physical properties: A possible model of analogical reasoning. *Psychological Record, 51*, 381–408.

Underwood, B. J., & Schulz, R. W. (1960). *Meaningfulness and verbal learning*. Chicago: Lippincott.

Vosniadou, S., & Ortony, A. (1989). Similarity and analogical reasoning: A synthesis. In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 1–18). Cambridge: Cambridge University Press.

Wong, E. (1993). Understanding the generative capacity of analogies as a tool for explanation. *Journal of Research in Science Teaching, 30*, 1259–1272.

Received: April 25, 2007
Final Acceptance: August 25, 2008