



ORIGINAL ARTICLE

# Testing and Training Analogical Relational Responding in Children With and Without Autism

Elle B. Kirsten<sup>1</sup>  · Ian Stewart<sup>1</sup> · John McElwee<sup>2</sup>

Accepted: 5 October 2021 / Published online: 10 January 2022  
© Association for Behavior Analysis International 2022

## Abstract

Analogical responding is ubiquitous in everyday language and cognition and is a key component in learning. However, there is relatively little behavioral research on analogical responding in young children. Kirsten et al. (2021) recently tested and trained an RFT-based model of analogy as the derived relating of relations in 5-year-old children using a relational evaluation procedure (REP) and the present study aimed to extend this work. In Experiment 1, two 5-year-old typically developing children were assessed and trained in relating both derived sameness and difference relations, thus extending the Kirsten et al. study, which had focused on difference relations alone. Following multiple exemplar training, correct responding increased to criterion levels and generalization was observed in both participants. In Experiment 2, similar results were seen for two children with autism spectrum disorder. Results are discussed in relation to previous research as well as the future testing and training potential of the REP with young children.

**Keywords** Analogy · Relational frame theory · Multiple exemplar training · Children · Autism · Relational evaluation procedure

Analogy, which involves identifying differing objects or events as being similar with respect to particular properties or features, is pervasive in daily language, social communication, and cognition (Bartha, 2019; Gentner et al., 2001; Hofstadter & Sander, 2013; Stewart et al., 2020). It is also critically important in various key domains of human activity including science, technology, and education (Bassok, 2001; Bod, 2009; Gentner, 1983; Hofstadter, 2001; Holyoak & Thagard, 1995; Morsanyi & Holyoak, 2010; Polya, 1945/2004; Sternberg, 1977; Stewart et al., 2004, 2013). As such, it is frequently used as a metric of intelligent behavior (Sternberg, 1977) and as a measure to predict academic success, for example, in the Law School Admissions Test (LSAT; Lapiana, 2004).

Considering the importance of analogy for intellectual development, cognitive-developmental psychologists have examined the emergence of this skill in typically developing young children. Early researchers believed that analogical

reasoning developed at the age of 12 or later and that children younger than this relied on simple associative strategies to solve analogies (Levinson & Carpenter, 1974; Lunzer, 1965; Piaget et al., 1977/2001; Sternberg & Nigro, 1980). In more recent research it has been argued that children as young as 3 can show analogy (Goswami & Brown, 1989, 1990; Richland et al., 2006). For example, Goswami and Brown (1989) purported to show that 3-year-olds could complete analogical tasks based on relations of causality (e.g., ice: melting ice) if they had knowledge of the domains involved (see also Alexander et al., 1989; Goswami, 1989). However, in a subsequent replication of Goswami and Brown, Rattermann and Gentner (1998) found that children younger than 5 were completing such tasks based on simple matching rather than analogical relations.

Most of the research on analogy has been conducted by cognitive psychologists. However, during the last 25 years behavioral psychologists have also begun to research analogy. For example, relational frame theory (RFT) researchers have explicitly recognized the theoretical importance of analogy (see e.g., Stewart et al., 2009). RFT is a contemporary contextual behavioral theory that proposes that arbitrarily applicable relational responding (AARR), or relational framing, is the key functional process involved in human

---

✉ Elle B. Kirsten  
e.kirsten2@nuigalway.ie

<sup>1</sup> National University of Ireland Galway, Galway, Ireland

<sup>2</sup> Wilkes Barre, USA

language and cognition (Dymond & Roche, 2013; Fryling et al., 2020; Hayes et al., 2001; Rehfeldt & Barnes-Holmes, 2009; Zettle et al., 2016).

RFT defines relational framing as a learned pattern of contextually controlled relational responding involving the three properties of mutual entailment, combinatorial entailment, and transformation of function. Mutual entailment (ME) involves deriving the reversal of a previously acquired unidirectional relation; for example, if taught that a novel coin A is worth more than a novel coin B then I might derive that B is worth less than A. Combinatorial entailment (CE) involves deriving the combination of previously acquired unidirectional relations; for example, if taught that A is more than B and B is more than C then I might derive that A is more than C and C is less than A. Transformation of function (TOF) involves deriving a novel function for a stimulus based on its derived relation with a second stimulus; for example, if I derive that coin B is less than coin A and B has a reinforcing function then I might subsequently respond to A as more reinforcing than B without additional training. From an RFT point of view, relational framing is the key process in human language and these properties of relational framing are what facilitate the generative properties of human language. Furthermore, there is by now substantial empirical evidence in favor of this thesis (see, e.g., Barnes-Holmes et al., 2020; Kirsten & Stewart, 2021; Stewart et al., 2013; Stewart & Roche, 2013).

Working within an RFT framework, Barnes et al. (1997) provided the first functional analytic definition of analogy as the derivation of a sameness or equivalence relation between derived relations. For instance, consider the analogy *diamond* is to *sapphire* as *rose* is to *orchid*. In this case, *diamond* and *sapphire* participate in an equivalence relation in the context of gemstones, and *rose* and *orchid* participate in an equivalence relation in the context of flowers. Thus, because these are both equivalence relations, we can derive a relation of equivalence between the relations themselves.

Barnes et al.'s (1997) work on analogy has been extended in a number of studies (e.g., Barnes-Holmes et al., 2005; Carpentier et al., 2002, 2003, 2004; Ruiz & Luciano, 2015; Stewart et al., 2001). One research thread relevant to the present article has focused on analogical responding in young children (e.g., Carpentier et al., 2002, 2003). For example, Carpentier et al. (2002) found that 9-year-olds and adults readily showed equivalence–equivalence (i.e., as in Barnes et al., 1997), but 5-year-old children failed to do so without supplementary prompting. In particular, the 5-year-olds required additional pretesting with compound–compound matching tasks involving trained (as opposed to derived) relations (e.g., A1B1–A3B3 and A1B2–A1B3) in order to successfully pass the derived compound relations (BC–BC) test. Carpentier et al. (2003) extended this work by examining if additional compound–compound testing would

also facilitate equivalence–equivalence performance in the absence of prior equivalence tests as had been seen in older participants. However, only 2 of the 18 five-year-old participants passed even with the additional compound–compound testing.

The studies just discussed all employed match-to-sample (MTS) procedures to train and test for both equivalence and equivalence–equivalence relations. One disadvantage of MTS is that it requires extensive baseline training before any testing or training of the critical relations can begin. For example, in Experiment 1 of Carpentier et al. (2002), the 5-year-old participants required an average of 234 baseline trials before testing could start. Furthermore, even after such extensive training, the relational network available for testing of derived relations or of training the capacity for derived relations if absent was severely limited. Although MTS is often used in studies of derived relations, alternative testing and training procedures may offer advantages in these respects, especially when examining relatively complex repertoires such as analogy or when working with younger children, or children with behavioral, developmental, or intellectual concerns, for whom training of deficient repertoires of derived relational responding may be especially important.

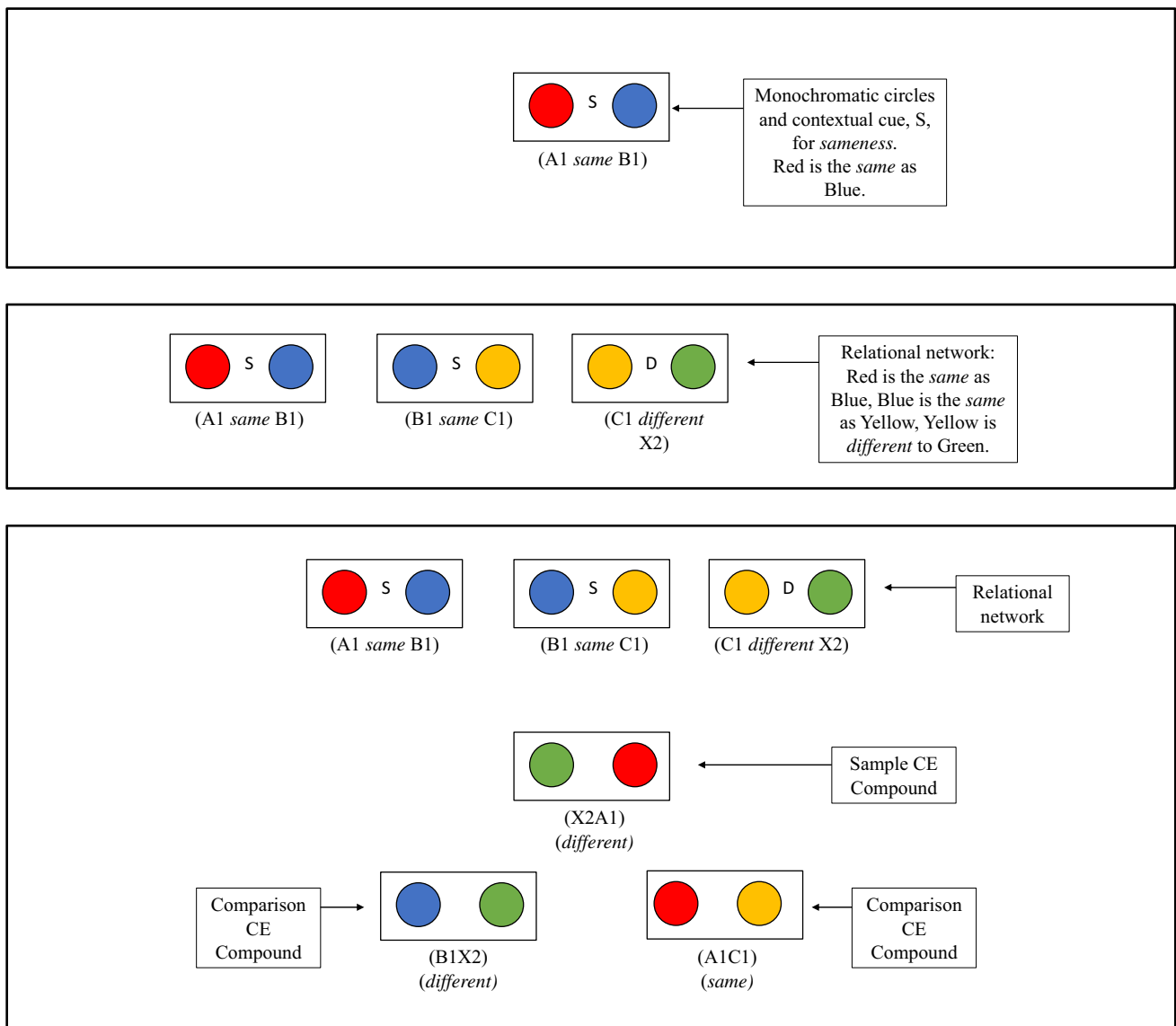
The RFT-based relational evaluation procedure (REP; see Barnes-Holmes et al., 2001; Stewart et al., 2004) offers one potential alternative to MTS. In the REP participants are required to evaluate or report on relational networks based on the presentation of contextual cues juxtaposed with relevant stimuli. For example, in Stewart et al. (2004), which used the REP to model analogy in adults, arbitrary shapes were first established as cues for “same,” “different,” “yes,” and “no.” Thereafter these cues were used to (1) establish relations of sameness and difference amongst arbitrary nonsense syllables and (2) to show that participants would evaluate analogical relationships involving these nonsense syllables coherently. For example, participants were shown to choose the “yes” cue when presented with the “same” cue juxtaposed with nonsense syllables in a relation of similarity and to choose the “no” cue when presented with the “different” cue juxtaposed with such a relation. The advantage that this procedure afforded over MTS was that, once the cues had been established, a completely novel set of nonsense syllables, and thus a completely new analogical relational network, could be presented on every trial, obviating the need for lengthy prerequisite training with respect to each set of nonsense syllables as would be needed with MTS.

The REP has been successfully utilized in several recent RFT-based studies to train relational framing in young children. For example, Cassidy et al. (2011) designed an REP-based automated AARR assessment and training program (see also Cassidy et al., 2016; Hayes & Stewart, 2016). The automated program presented multiple exemplars of relational statements involving nonsense words juxtaposed with

contextual cues (e.g., “CUG is the SAME as DAX,” “DAX is the SAME as YIM”), followed by questions requiring relational derivation based on those statements (e.g., “Is DAX the SAME as CUG?,” “Is CUG the same as YIM?”). Cassidy et al. successfully trained key patterns of relational framing in 8- to 12-year-old children and saw significant boosts in their intellectual performance, thus suggesting the potential utility of the REP format in training relational framing in children.

Kirsten and Stewart (2021) designed a relatively comprehensive REP-based relational assessment to test a variety of relational frames across four levels of responding including nonarbitrary relations, nonarbitrary analogical relations, arbitrary relations, and arbitrary analogical relations

in young children, including children not yet able to read. The researchers taught the children to respond to relational networks composed not of textual stimuli but instead of colored circles as the relata juxtaposed with single letters as contextual cues (e.g., S for *sameness*, D for *difference*). For example, children were taught that given a red circle and a blue circle separated by the contextual cue “S” they should subsequently treat the red and blue circles as the same or equivalent (i.e., “Red is the same as Blue”; see Fig. 1, top panel, for an illustration of the stimuli). For testing analogy, compound stimuli (i.e., one sample compound and two comparison compounds; see Fig. 1, bottom panel, for an example) composed of colored circles in either same or difference relations were presented below a relational network



**Fig. 1** Format for Presenting Relational Stimuli. *Note.* First panel: monochromatic circles plus contextual cue, S; second panel: relational network; third panel: relational network plus analogical stimuli.

(see Fig. 1, middle panel, for an example) and children were required to choose same with same and difference with difference relations. This format allowed young children, including nonreaders, to report on and evaluate multiple exemplars of arbitrarily applicable relational networks defined by specifically selected contextual cues.

In a more recent study, Kirsten et al. (2021) adapted the arbitrarily applicable relational stages of the REP-based assessment in Kirsten and Stewart (2021) to test and train analogical relations in 5-year-old children. Kirsten et al. found that after direct training in relating combinatorially entailed relations using the REP, all participants demonstrated analogical responding across multiple stimulus sets without requiring additional prompting. However, one potential issue in Kirsten et al. was that the relational networks across all of the stimulus sets permitted testing of only combinatorially derived *difference* relations. This was because the relational networks included only four arbitrary stimuli and three direct relations—two sameness and one difference relation (e.g., Red is the same as Blue, Blue is the same as Yellow, Yellow is different to Green; refer to the bottom panel in Fig. 1 for an illustrative example). This relatively curtailed network permitted only one combinatorially derived sameness relation (in the case of the example above, Red : Yellow is the only possible combinatorially derived sameness relation) per trial and hence, there was no opportunity to test participants for the matching of two combinatorially entailed sameness compounds. In contrast, Carpentier et al. (2002, 2003) had trained and tested for both combinatorially derived sameness and difference relations. Kirsten et al. therefore suggested that in future research in this domain the array of stimuli in the relational network should be increased in order to allow for both combinatorially derived sameness and difference relations.

In Experiment 1 of the present study, we sought to extend Kirsten et al. (2021) by modifying the REP training to include a larger array of stimuli, thus permitting the testing of both combinatorially entailed sameness and difference relations. One other methodological difference was that instead of employing multiple exemplars of the relation of derived relations in the training intervention, we employed multiple exemplars requiring the relation of directly presented relations. This was in order to examine, analogous to Carpentier et al., whether inducing children to engage in the relation of directly presented relations might prompt them to subsequently show the relation of derived relations.

Experiment 2 of the present study was a replication of Experiment 1, but participants were children with autism spectrum disorder (ASD). Characterized by impairments in social interaction and social communication (American Psychiatric Association, 2013), ASD currently affects 1 in 54 children in the United States (Maenner et al., 2020). It has been argued that children with ASD face significant

language comprehension challenges due in part to their difficulty in understanding figurative language (Kalandadze et al., 2018; Persicke et al., 2012). However, the acquisition of analogical language in children struggling with ASD has received little attention. In the only extant behavioral study in this area, Persicke et al. successfully taught metaphorical language to three participants with ASD using multiple exemplar training. In addition, Persicke et al. found that participant responses generalized to untrained, novel metaphors. However, two notable experimental limitations were observed: participant history with the metaphors could not be controlled, and the relative difficulty of the metaphors was not quantified and thus difficulty across metaphors could not be established. In the present study, in contrast, all relations were established among arbitrary stimuli within the experimental task thus obviating the need to control for task variance and participant history with language.

## Experiment 1

### Method

#### Participants and Setting

Participants were typically developing children enrolled in a private elementary school on the East Coast of the United States. Six potential participants were given preassessment testing; of those, three passed the preassessment and two of those proceeded to baseline sessions. P1.1 was a female aged 6 years and P1.2 was a female aged 5.25 years. Participants were selected for inclusion based on their performance on an adaptation of the relational assessment used in Kirsten and Stewart (2021). All probe and training sessions were administered by the researcher in an otherwise unoccupied classroom of participants' school during school hours. The researcher sat next to the participant at a standard school desk. A second, independent observer sat approximately 3 feet away from the desk on the other side of the participant with a full view of the participant and the computer screen.

Ethical approval for recruitment of participants was obtained from the research ethics committee of the lead researcher's host institution. Consent for conducting the study was also obtained from the principal of the school. Caregiver consent was obtained for each child who participated, and verbal assent was obtained from each of the participants.

#### Experimental Design

A multiple baseline design across participants was used in this study. Details for each condition of the multiple

baseline are described below. In order for the second participant to enter the training condition, the first participant had to meet the probe criterion (i.e., scoring at least five out of six (83%) correct on the probe trials).

### Materials and Apparatus

A 13-in MacBook running Microsoft PowerPoint was used to present trials. Individual stimuli included colored circles (either 0.5 in or 1 in in diameter, depending on trial type) and letters “S” for *sameness* and “D” for *difference* relations in Calibri or Arial, size 24 or 36 fonts (depending on the trial type; see first panel in Fig. 2). During the relational section of the preassessment, trials included an array of such stimuli at the center of the screen that were designated by the experimenter as participating in a relational network (see second panel in Fig. 2). During the analogical section of the preassessment, as well as during study probe and training sessions, similar to the relational section, trials included an array of stimuli that were designated as participating in one of four possible relational networks. In this section, however, these appeared in the left portion of the screen only. In the right portion of the screen there appeared either (1) a sample compound element for the pre-analogy relational trials or, (2) a sample compound and two comparison compound stimuli below the sample on the bottom left and right of the screen, separated by a black line for analogy trials (see third panel in Fig. 2 for an illustrative example of relational networks and compound elements).

The relational networks in the left portion of the screen included 16 monochromatic circles and the relational cue, S for *same*, to delineate relations between particular circles. For example, one possible array might be represented as follows: [Black Circle] [S] [Gray Circle], [Yellow Circle] [S] [White Circle], [White Circle] [S] [Blue Circle], [Green Circle] [S] [Orange Circle], [Turquoise Circle] [S] [Green Circle], [Pink Circle] [S] [Brown Circle], [Red Circle] [S] [Brown Circle], [Purple Circle] [S] [Gray Circle]. This array might allow a potential participant to derive four equivalence relations including black, gray, and purple; yellow, white, and blue; green, orange, and turquoise; and pink, brown, and red. The compound elements that appeared on the right in black, outlined rectangles as the sample and comparison stimuli were each composed of two of the monochromatic circles from the relational network but did not contain relational cues (see top section of third panel in Fig. 2). For example, one such compound might be designated as [Yellow Circle][Red Circle]. Each slide had four relational networks within the array of 16 stimuli and one set of task stimuli.

## Procedure

### Overview

The following will provide procedural details of the relational preassessment and the multiple baseline across participants design of the study. The preassessment was administered first in one session and took 5–20 min to complete. Following preassessment, participants entered the baseline condition of the multiple baseline design, followed by a brief pretraining probe, followed by the intervention condition, which included training and probe sessions. Table 1 provides a schematic overview of procedures. An average of eight sets were run per day during the intervention condition; a probe set took on average 3–4 min to complete, and a training set took approximately 1–2 min to complete. Both participants started baseline sessions at the same time, and both participants took approximately 3 weeks to complete the study once the training condition was implemented.

### Preassessment

Tasks adapted from the arbitrary relational and analogical stages in Kirsten and Stewart (2021) were used in the relational preassessment (see Fig. 3). The preassessment, which comprised 82 trials in total, included four main sections outlined in the following.

**Section 1: Relational tasks** The first six trials in the preassessment introduced the participant to the contextual cues themselves (i.e., S and D; refer to Fig. 3). Next, the researcher introduced the participant to simple, arbitrary relational networks. The participant was shown a computer screen displaying a relational network, for example: [Red Circle] [S] [Blue Circle] (see Fig. 3). The assessor instructed the participant to look at the screen and said, “Let’s read this: Red is the same as Blue” (in this example and hereafter, *reading* refers to vocally identifying the stimuli and relational cues in the relational network in sequence from left to right, similar to textual reading). After delivering the instruction, the assessor asked yes/no and same/different questions about the relational networks, including questions about directly presented relations (e.g., “Is Red the same as Blue?” or “Is Red the same or different to Blue?”), and questions requiring reversal of the directly presented relation (i.e., mutually entailed relations such as “Is Blue the same as Red?”).

The next set of trials included more than two stimuli, and questions became increasingly difficult and required responding not only to directly presented (DP) and mutually entailed (ME) type questions but also to questions that required combinatorial entailment (CE) of directly presented relations (see Section 1 of Fig. 3). The first set of

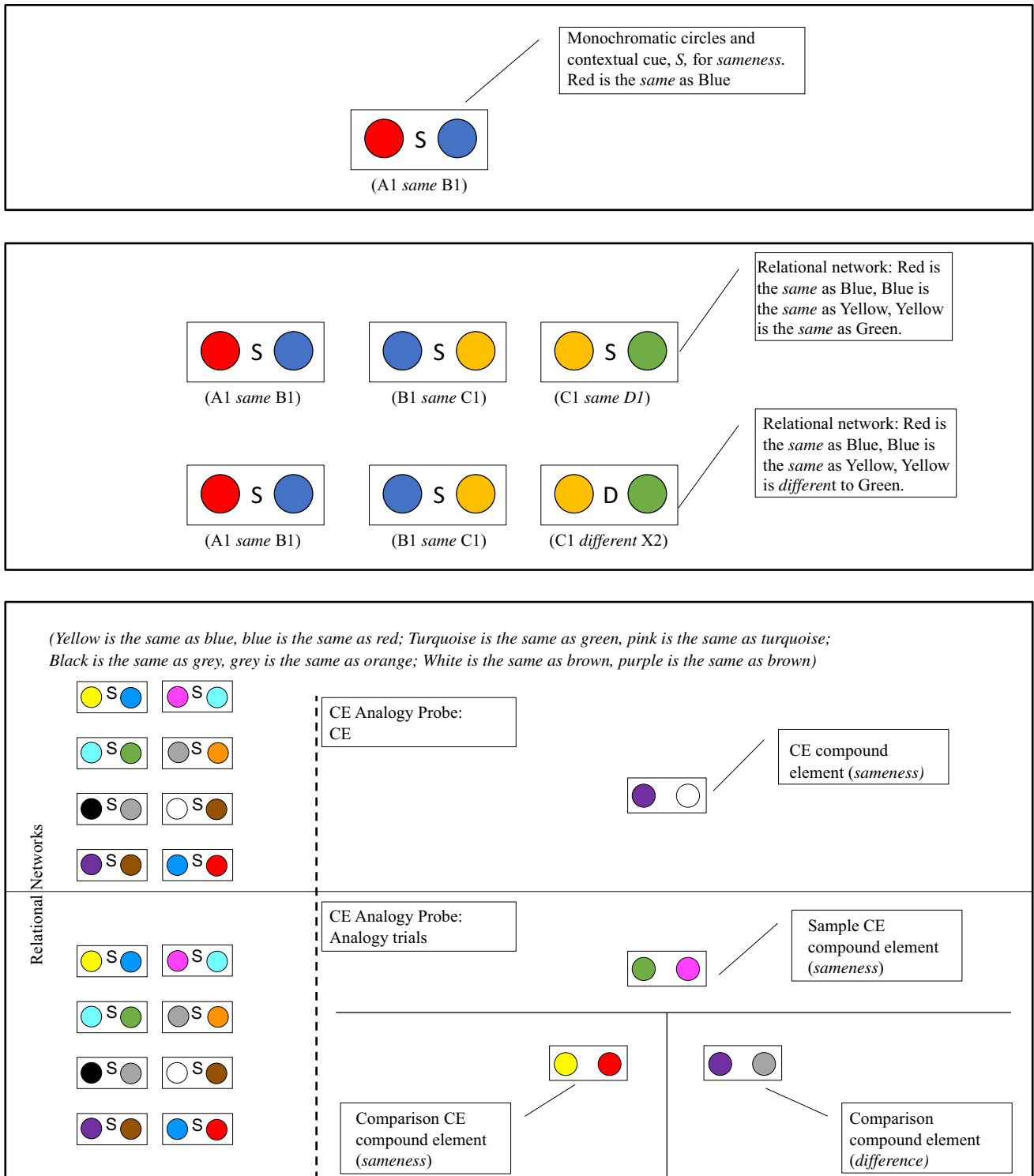


Fig. 2 Preassessment Screening Tool Stimuli Arrangement

questions in this section referred to a relational network in which three *sameness* relations were presented; the second set referred to a relational network including two *sameness* cues and one *difference* cue. An example of the latter set might be as follows: The relational network [Red Circle]

[S] [Blue Circle], [Blue Circle] [S] [Yellow Circle], [Yellow Circle] [D] [Green Circle] is presented on the screen followed by questions regarding combinatorially entailed relations among the stimuli (e.g., “Is Red the same/different to Yellow?”).

**Table 1** Schematic presentation of preassessment and multiple baseline conditions

Pre-Assessment		
Phase	Relational Targets	Passing Criteria %
	1. Identifying contextual cues	100
	2. Reading relational networks	100
	3. Testing DP and ME relations	100
	4. Testing CE relations	80
Inclusion Criteria	5. Directly presented compound trials	100
	6. Directly presented analogy trials	NA
	7. CE sorting task	100
	8. CE trials with sorting	80
	9. CE trials w/o sorting	80
Multiple Baseline Study		
<u>Baseline condition</u>		
	CE Probe Set 1: 6 CE trials + 6 CE analogy trials	100
	Pre-Training DP Analogy Probe: 6 trials	100
<u>Intervention</u>		
	DPA-Training Phase 1: 6 trials	100
	DPA+XF Training Phase 2: 6 trials	100
<u>Probes</u>		
	1. CE Probe Set 1: 6 CE trials + 6 CE analogy trials	100
	2. Novel CE Probe: 6 CE trials + 6 CE analogy trials	100
	3. Novel CE Probe w/ distractor: 6 CE trials + 6 CE analogy trials	100

Note. *DPA* directly presented relations, *ME* mutually entailed relations, *CE* combinatorially entailed relations

**Section 2: Directly presented compound elements** The next set of tasks presented compound elements including a stimulus composed of two, side-by-side monochromatic circles identical to the circles in the relational network, without the relational cue, S (see Section 2 in Fig. 3). For example, one such compound might be designated as [Red Circle][Yellow Circle]. In each trial, the relational networks were presented at the left of the screen. To the right of the network, a white rectangle with a black outline contained the compound element (i.e., two differently colored circles identical to two of the circles in the relational network). The researcher and participants read the compound element together and then the researcher said, “Look here [points to relational networks], to figure out if these [points to the element compound] are

the same or different. Remember to look here [points to relational network] to help you figure it out.”

All the compound elements in this section were directly related (*sameness* relation) or not in the same network (*difference* relation). For example, the relational networks [Black Circle] [S] [Gray Circle], [Yellow Circle] [S] [White Circle], [White Circle] [S] [Blue Circle], [Green Circle] [S] [Orange Circle], [Turquoise Circle] [S] [Green Circle], [Pink Circle] [S] [Brown Circle], [Red Circle] [S] [Brown Circle], [Purple Circle] [S] [Gray Circle] are presented at the left of the screen, and the compound stimulus (e.g., [Pink Circle] [Gray Circle]) is presented to the right of the network; thus, the participant might look at the relational networks and find that pink and gray are the same (the compound element)

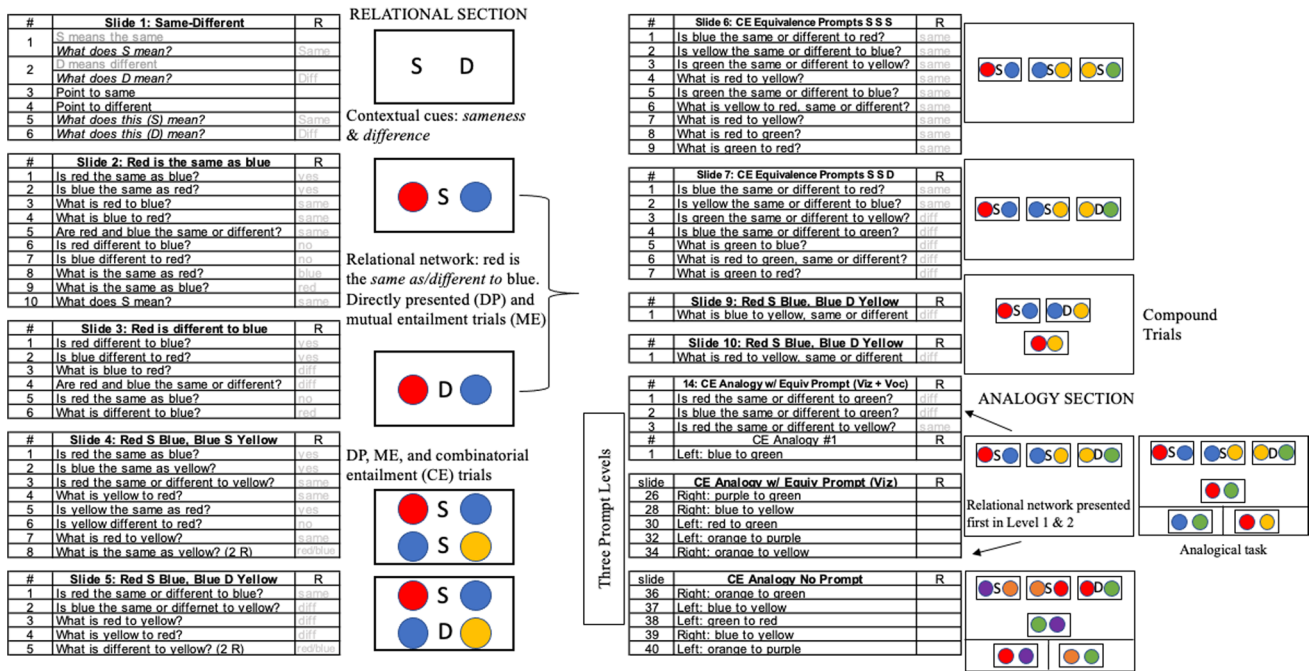


Fig. 3 Preassessment Screen Tool. Note. Adapted from Kirsten and Stewart (2021)

because the relation is directly presented as [Purple Circle] [S] [Gray Circle]. Each slide included the four relational networks and one compound stimulus.

Following the compound questions, the second task in Section 2 presented directly presented analogical stimuli (see Section 2 in Fig. 3 and the first panel in Fig. 4). The directly presented analogical stimuli were presented to the right of the relational networks. For example, the relational networks [Black Circle] [S] [Gray Circle], [Yellow Circle] [S] [White Circle], [White Circle] [S] [Blue Circle], [Green Circle] [S] [Orange Circle], [Turquoise Circle] [S] [Green Circle], [Pink Circle] [S] [Brown Circle], [Red Circle] [S] [Brown Circle], [Purple Circle] [S] [Gray Circle] are presented at the left of the screen, and the directly related compound sample element (e.g., [White Circle][Blue Circle]) is presented to the right of the relational networks, and the two comparison compound elements (e.g., directly related [Purple Circle][Gray Circle] and not related [Pink Circle] [Black Circle]) are presented below the sample. On each trial, the researcher delivered the instruction, “Look at this one at the top [pointing to the sample compound]. Which one of these [pointing to each of the comparison compounds in turn] is like this one at the top?” (see first panel in Fig. 4). The participant had to refer to the relational networks to determine if the stimuli within the compound elements were the same or different.

**Section 3: Combinatorial entailment** In the first task in this section, participants were given 12 monochromatic tokens

that matched the colors of the circles in the relational networks, and a sheet of paper divided equally into four sections (see second panel of Fig. 4). One token from each relational network was placed in its own section on the paper. The researcher gave the instruction, “Look here [points to the relational networks] to figure out which circles go with each other. There are four sets of circles and three circles in each set.” Sorting responses were scored as correct or incorrect for a total of 12 responses.

The next task in Section 3 required the same sorting task followed by six questions regarding the combinatorially entailed relations among the stimuli. A Power-Point slide was presented showing a combinatorially entailed compound element and the instruction, “Do these circles go together? Look here [point to the relational network] and here [point to the four sets of tokens in front of them] to figure it out” (see third panel in Fig. 4).

The final task in Section 3 tested for combinatorial entailment without the tokens. As in the previous task, participants were shown a screen with the relational network on the left and a compound element to the right of it and given the instruction, “Do these circles go together? Look here [point to the relational network] to figure it out.” If participants scored below 80%, they were instructed to use the tokens again and all trials were re-presented. Following the token trials, the token-free trials were readministered. Potential participants had to score at least 80% correct to proceed to Section 4.



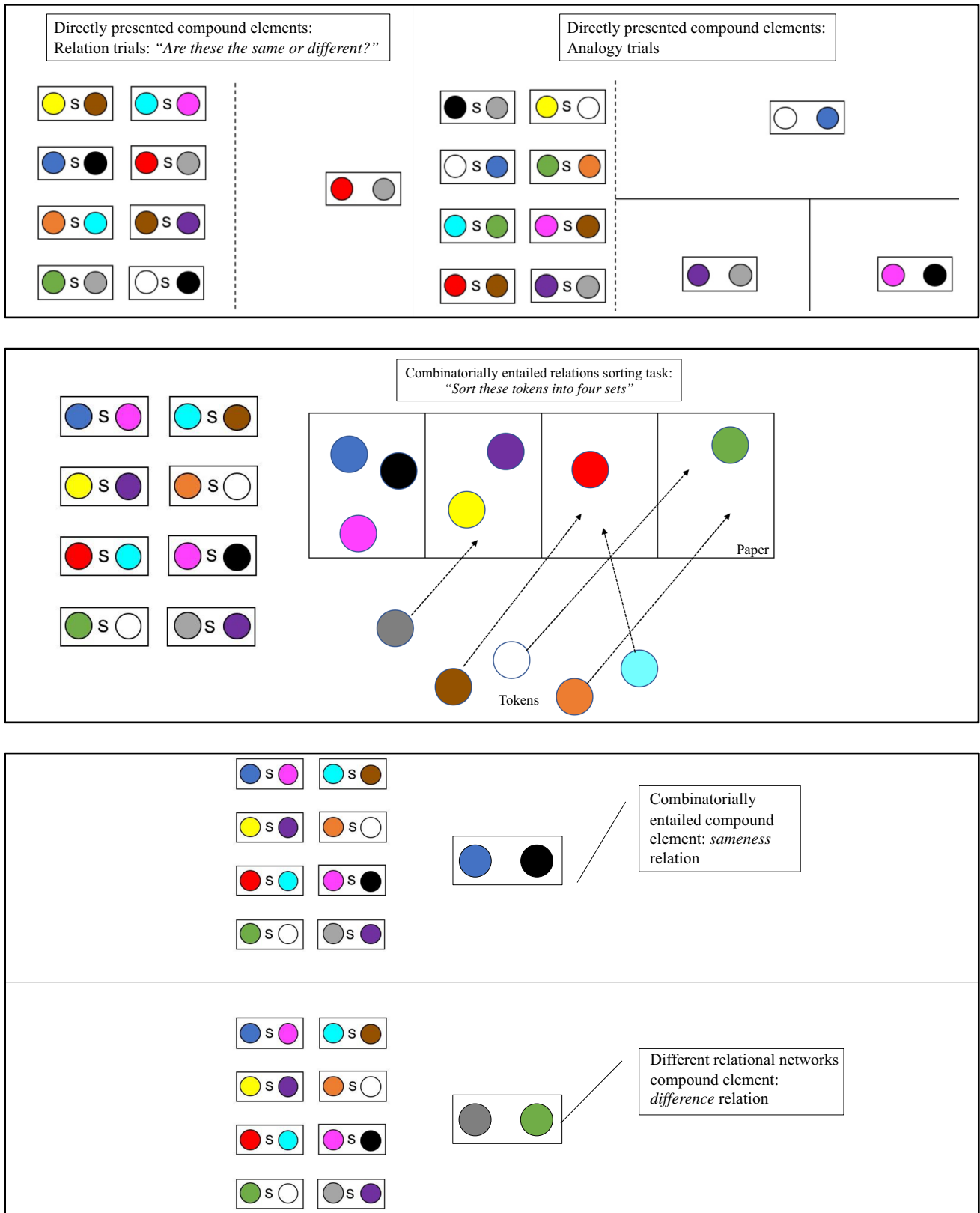


Fig. 4 Preassessment Section 2: Directly Presented Analogy; Section 3: Combinatorial Entailment

**Section 4: Relating combinatorially entailed relations: Analogical relations** There were two tasks in the analogical section of the preassessment. The first task was identical to the last task just described in Section 3. Six combinatorial entailment (CE) trials were presented. Following the six CE trials, the same relational network was presented with six analogy trials. The analogical stimuli included the four relational networks and three compound elements composed of two circles (i.e., the sample and two comparisons; see Fig. 5). On each trial, the researcher delivered the instruction, “Look at this one at the top [pointing to the sample compound]. Which one of these [pointing to each of the comparison compounds in turn] is like this one at the top? Look here [points to relational networks] to help you figure it out.” In order to proceed to the baseline condition, participants had to score at least five out six correct (83%) on the CE relational trials and fail the analogy trials.

**Multiple Baseline Conditions**

The multiple baseline design across participants comprised a baseline condition including unreinforced baseline sessions, a brief, two-session pretraining probe condition in which directly presented analogical relations were assessed, and the intervention condition in which training and multiple probe sessions were administered. The study included two types of probe trials: Combinatorially Entailed Analogy Probes (CE

Probes) and Combinatorially Entailed Analogy Probes with a Distractor (CE+D Probes) (see Fig. 6 for an illustrative example of each probe type). CE Probe Set 1 was administered in all baseline sessions, and it was the first probe set after training commenced, followed by novel CE Probes, and CE+D Probes, in that order.

The CE Probe sets were identical to the CE relational and analogy trials in the preassessment (Section 4 of the preassessment). Comparison compounds never included either stimulus presented in the sample compound. In CE+D Probes, one of the comparison compounds included one of the stimuli from the sample compound. All sample and comparison compound elements were comprised of either combinatorially entailed sameness relations, or relations of difference in which the stimuli did not belong to the same relational network. Both CE and CE+D Analogy Probes included six CE relational trials and six CE analogy trials as described in Section 4 of the preassessment.

During the six CE relational trials in both probe types (i.e., CE and CE+D), the participant and researcher looked at the laptop screen with the relational network on the left and a compound element to the right of it, and the researcher asked, “Do these circles go together? Look here [point to the relational network] to figure it out.” No feedback was provided for correct or incorrect responding.

During the six CE analogy trials in both probe types, the participant and researcher looked at the laptop screen, the researcher

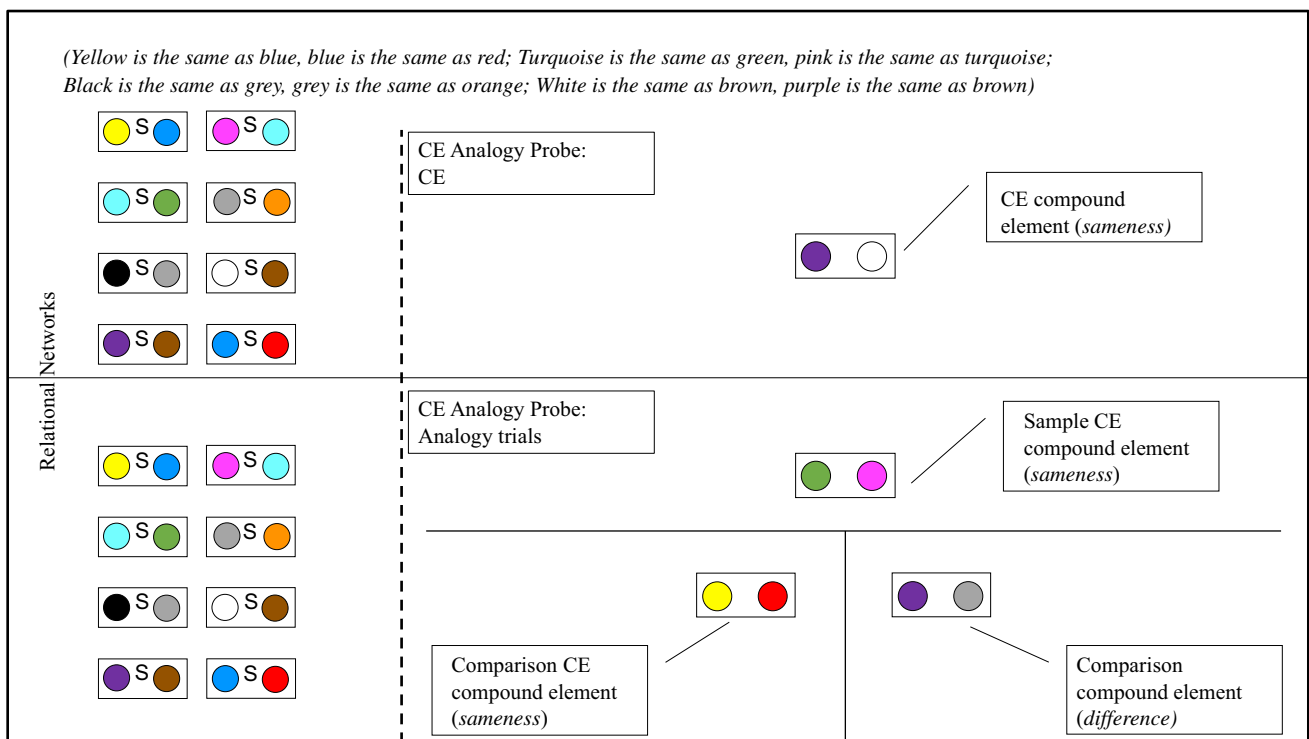
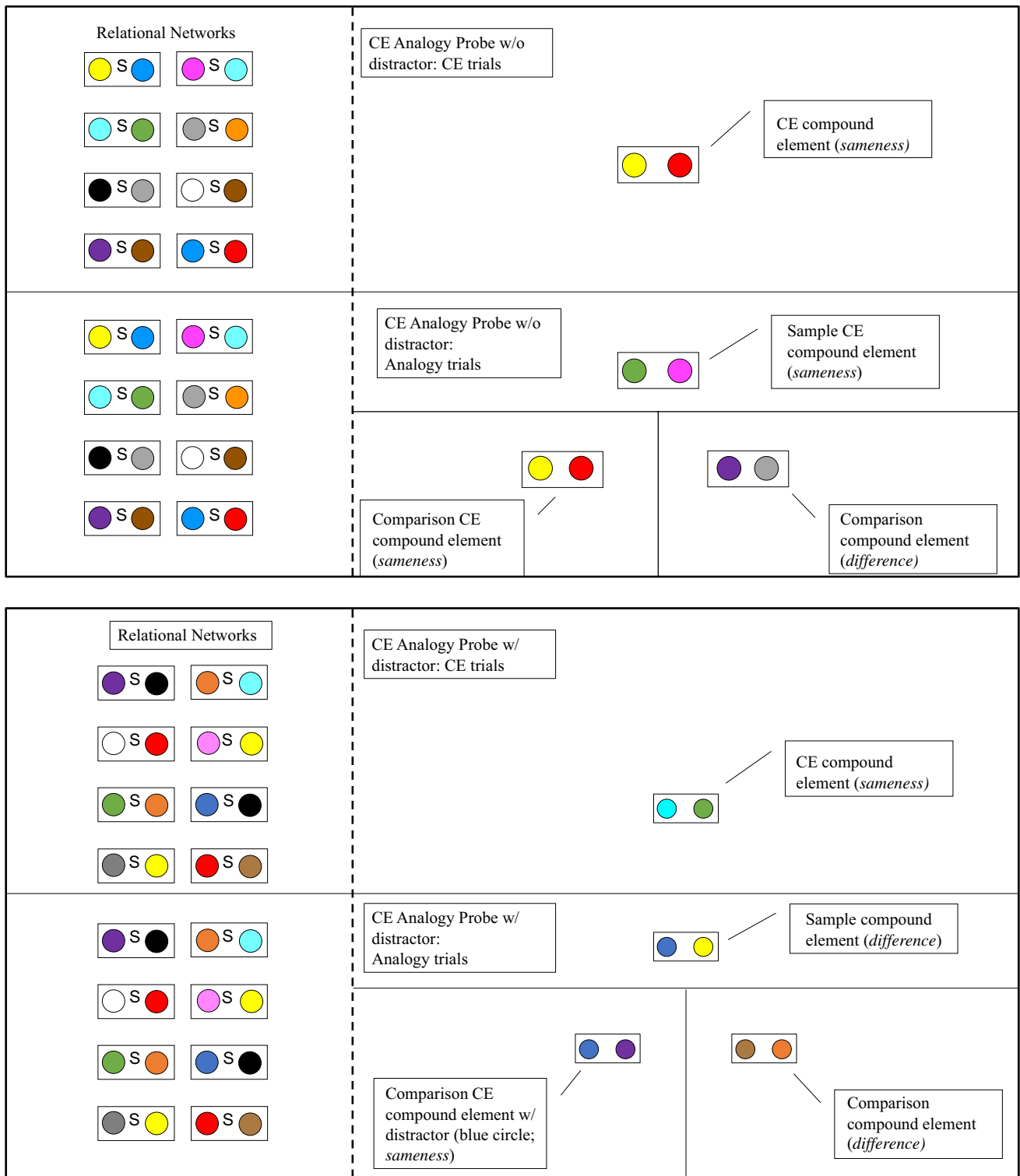


Fig. 5 Preassessment Section 4: Analogical Relations



**Fig. 6** Two Probe Types: CE Analogy and CE+D Analogy

instructed the participant to look at the sample compound element and said, “Look at this one at the top [points to the sample compound], which one of these [points to comparison compounds] is like this one at the top?” No feedback was provided for

correct or incorrect responding. The same relational network was used across all trials within a probe set. Passing criteria required responding correctly on all six trials (100%) the first time the probe was presented, or scoring 100% correct twice consecutively.

Following baseline CE Probe Set 1 sessions, a pretraining probe condition was implemented in order to assess participant responding to directly presented analogical relations (DPA-Probe). The stimulus format was the same as in the CE Probes except all stimuli in the compound elements in the pretraining probe condition were directly related (a *sameness* relation) or not in the same network (a *difference* relation). During the DPA-Probe, the researcher instructed the participant to look at the screen and said, “Which one of these [points to comparisons] is like this one at the top [points to sample]? If the participant responded correctly to all six trials in the DPA-Probe twice consecutively, CE Probe Set 1 was re-presented. If the participant did not pass the DPA-Probe it was probed a second time. If the participant failed again, two trials demonstrating directly presented analogical responding were presented. During the demonstration, the presenter said, “Look here [points to sample], now point to this one [points to correct comparison], this one goes with this one [points to sample], you do it.” There were only two DPA-Probe trials. Passing criteria for all Probes was five out of six correct (83%).

The training condition was implemented following the two DPA pretraining probe sessions. The training condition included two phases—Phase 1: Directly Presented Analogy Training (DPA-Training) and Phase 2: Directly Presented Analogy Training Plus Extra Feedback (DPA+XF Training). A modified version of the Greer and Ross (2008) decision-making protocol was followed during the training condition. If Phase 1 training data showed five ascending data paths, then training in Phase 1 continued. If Phase 1 data showed five variable or five descending data paths, Phase 2 training would be implemented. Both participants required Phase 2 training.

Training sets included the same relational network format as in the CE and DPA Probes (see Fig. 7). However, all the stimuli in the compound elements were directly related as in the DPA Probes. Each training set included six directly presented analogy trials presented on six PowerPoint slides.

In Phase 1 DPA-Training, the participant was shown the analogical stimuli on the computer screen including the four relational networks on the left of the screen and directly presented compound elements to the right of the relational networks. In each trial, the relations between the circles in the sample were either directly presented in the relational network and therefore a *sameness* relation, or they were not in the same relational network and therefore a *difference* relation. Once the participant was looking at the screen the researcher gave the instruction, “Look at this one at the top [points to the sample compound], which one of these [points to comparison compounds] is like this one at the top? Remember to look at the information here on the side [points to relational networks] to help you figure it out.” In Phase 1 DPA-Training, the participant received yes/no feedback for correct or incorrect responding. A correct trial was consequted with, “Yes, that is correct!” and an incorrect trial was consequted with, “No, that is incorrect.” The following trial was presented regardless of correct or incorrect responses.

In Phase 2 DPA+XF Training, more instruction and feedback were included in each trial. The participant was shown the screen with the analogical stimuli and given the instruction, “Look at this one first [points to sample], and figure out if it’s the same or different. Now look at these here [points to comparisons], which one of these is like this one [points to sample]? Remember to look at the information here on the side [points to relational networks] to help you figure it out.” A correct trial was consequted with, “Correct/good/yes! They’re both the same/different.” Or “No, this one (points to sample) goes with this one (points to the correct comparison).”

Passing criterion was 100% correct once on the DPA-training trials. When the participant met criteria, the baseline probe, CE Probe Set 1, was readministered, including the six relational trials and the six analogy trials. If participants failed the CE probe, they went back into training and had

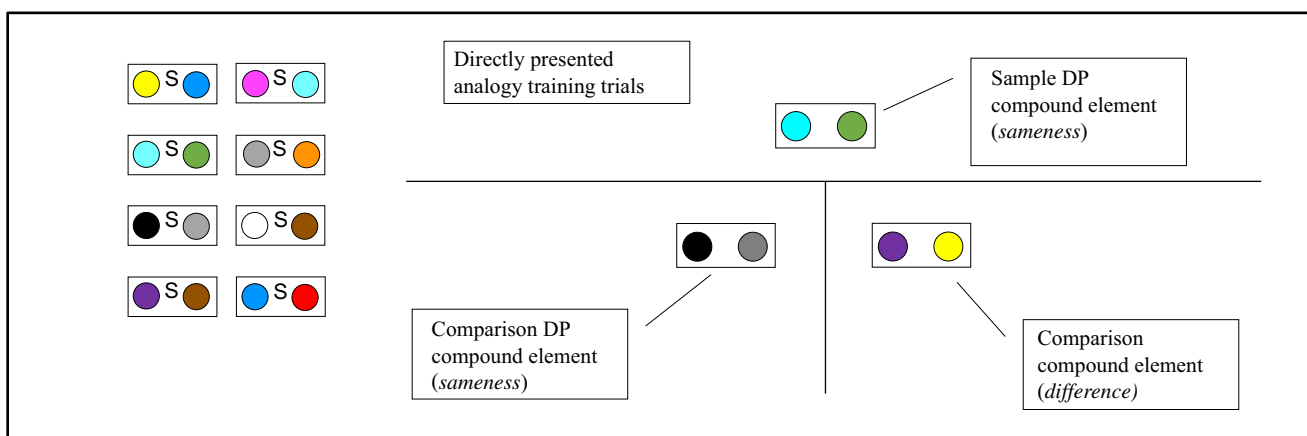


Fig. 7 Training Stimuli: Directly Presented Compound Elements for Directly Presented Analogical Responding

to score 100% on training trials before CE Probe Set 1 was re-presented. If the participant passed the six analogy trials at 100% correct, another CE Probe Set 1 was administered. If they passed again at 100% correct, a novel CE probe was administered. If they passed the novel CE probe, a CE+D probe was administered.

In summary, the study included a relational preassessment for screening potential participants; a baseline condition in which relating combinatorially entailed relations (CE analogy) was tested; a brief pretraining probe condition in which relating directly presented relations (DP analogy) was tested; and a training condition in which relating directly presented relations (DP analogy) was trained and CE analogy probe trials were presented.

### Interobserver Agreement

Procedural fidelity checks and interobserver agreement (IOA) were determined for baseline, probe, and training conditions by a trained research assistant. Procedural fidelity was assessed through the use of a fidelity checklist in which each trial in each condition was scored as either correct or incorrect; correct presentation required adherence to all relevant procedural criteria based on condition and trial type including presentation and use of the appropriate feedback. Procedural fidelity was assessed for 46% of all trials and was 98%. IOA was calculated on a trial-by-trial basis for each probe and training trial. IOA was assessed for 28% of Participant P1.1's sessions, IOA was 100%; and 30% of P1.2's sessions, IOA was 100%.

## Results and Discussion

### Overview

Following training, both participants successfully showed analogical responding during CE Probe sets, including the original CE Probe Set 1 used during baseline testing, a novel CE Probe Set 2, and the generalization probe, CE+D Probe (see (see Table 2 for a summary of condition names and acronyms). Both participants scored 100% correct on CE Probe Set 1. Participant P1.1 scored 100% correct on the novel CE probe and the CE+D generalization probe, and Participant P1.2 scored 83% correct on the novel CE probe and the generalization probe. Both participants required Phase 2 DPA+XF Training (see Fig. 8).

### Preassessment

As previously indicated, six potential participants were tested for relational responding on the preassessment.

**Table 2** List of condition names and acronyms

Probe and Training Conditions	Acronym
Combinatorially Entailed Probe	CE Probe
Combinatorially Entailed Probe w/ Distractor	CE+D Probe
Directly Presented Analogy Probe	DPA-Probe
Directly Presented Analogy Training (Phase 1)	DPA-Training
Directly Presented Analogy Training Plus Extra Feedback (Phase 2)	DPA+XF

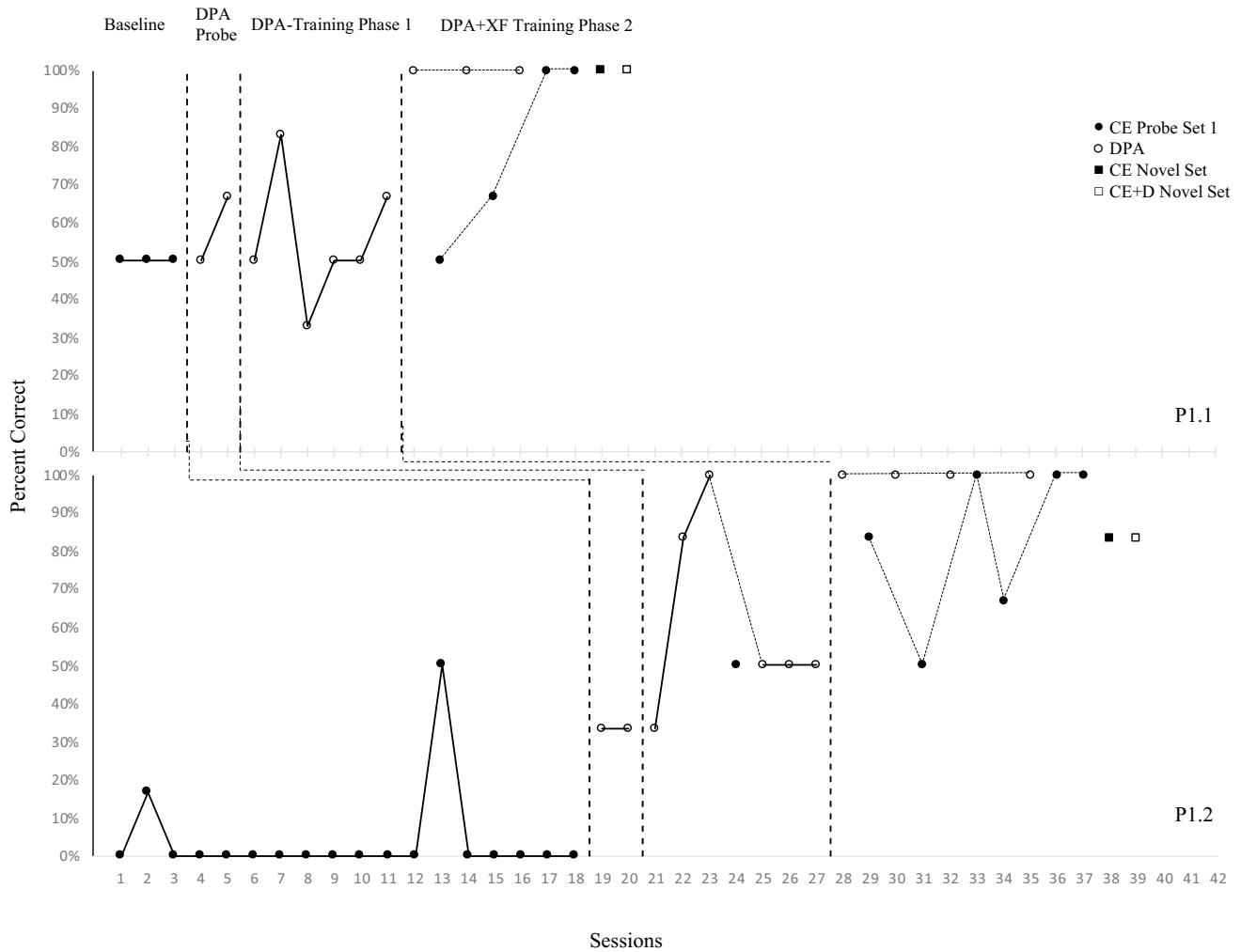
Three of these children passed the combinatorially entailed relational tasks and two of those proceeded to baseline sessions. Of the participants who met criteria on the preassessment, both P1.1 and P1.2 scored 67% correct on the first set of CE trials, and 100% correct on the second attempt. See Table 3 for preassessment scores.

### Participant 1.1

Participant P1.1 scored 50% correct on all baseline combinatorially entailed analogical relations sessions (CE probes), and she scored 50% and 67% on the pretraining Directly Presented Analogy Probe (DPA-Probes) sessions (see Fig. 8 for participant results). Participant P1.1 did not meet the training criteria during the Directly Presented Analogy Probe (DPA-Probes) pretraining probe sessions; thus, Phase 1 Directly Presented Analogy Training (DPA-Training) was implemented. Participant P1.1's training scores did not increase to passing levels after six training sessions (i.e., five data paths) and thus Phase 2 Directly Presented Analogy Training Plus Extra Feedback (DPA+XF Training) Training was implemented. Participant P1.1 scored 100% during the first DPA+XF Training session but her CE probe score stayed at baseline level (50%). Following two more DPA+XF Training sessions, P1.1's CE probe score increased to 100% correct for all probe sets including CE Probe Set 1, CE Probe Set 2 (novel probe), and the CE+Distractor probe.

### Participant 1.2

Participant P1.2 maintained low levels of responding during baseline CE Probe sessions, and she scored 33% correct on both pretraining DPA-Probe sessions. Participant P1.2 met the training passing criteria after three DPA-Training sessions but failed CE Probe Set 1. After scoring 50% thrice consecutively in DPA-Training sessions following the CE Probe, Phase 2 DPA+XF Training was implemented. P1.2 scored 100% during all DPA+XF Training sessions, and required three training sessions before scoring 100% correct on CE Probe Set 1. P1.2



**Fig. 8** Experiment 1: Participant Responding in Analogy Probe and Training Sessions. *Note.* CE Probe Set 1: Combinatorially entailed analogical responses; DP Probe: Directly presented analogy probes; DPA-Training Phase 1: Directly presented analogical training w/ min-

imal feedback; CE+D: Combinatorially entailed analogical responses with distractor; DPA+XF Training Phase 2: Directly presented analogical training w/ increased feedback

**Table 3** Experiment 1: Preassessment Relational and Analogical Scores (Percent Correct)

Participant	Sort Tokens	CE Sort w/ Tokens Relational	CE Sort w/o Token		CE & Analogy Probe	
			1st Attempt Relational	2nd Attempt Relational	Relational	Analogy
P1.1	100	83	67	100	83	50
P1.2	100	100	67	100	83	0

*Note.* Inclusion criteria required passing the relational test at 80% or above and failing the analogy test

required one more training DPA+XF Training session before meeting the passing criteria for CE Probe Set 1. P1.2 scored 83% correct on both CE Probe Set 2-novel and the CE+Distractor Probe.

After direct training in relating directly presented relations, both participants showed analogical responding according to RFT’s conception of analogy as the derived

relating of relations. Both participants required Phase 2 DPA+XF Training in order to meet passing criteria on both training and probe trials. DPA+XF Training included more instruction and feedback in each trial compared to the minimal instruction and feedback in the DPA-Training. Regarding Participant P1.2, it is possible that the extended time in baseline sessions affected her motivation to respond in the

training phase. She was not motivated to respond to the CE relation trials or the analogy trials in the baseline condition. Only after implementing DPA+XF Training did Participant P1.2's probe scores increase.

The results from the multiple baseline showed that the directly presented (relations) analogy (DPA) training procedure was an effective intervention for training analogy and eliciting generative CE analogical responding as shown by the generalization data. Both participants passed baseline CE Probe Set 1 as well as a novel CE Probe following DPA analogy training. Furthermore, correct analogical responding generalized to the CE+D Probe.

This extends Kirsten et al. (2021) who used this RFT approach to examine analogy in young children. Furthermore, these data support the Carpenter et al. and Kirsten et al. findings that 5-year-old children are capable of analogical responding.

## Experiment 2

Experiment 1 was replicated in Experiment 2, but participants were children diagnosed with ASD.

### Method

#### Participants and Setting

Three potential participants volunteered to take part but only two completed the study. Participants were two males with an independent ASD diagnosis for whom the first author of the present study provided 1:1 applied behavior analytic services. Participant P2.1 attended a private behavioral and learning center in New York City, 5 days a week for 5 hr per day. Participant P2.2 attended a private school that provided a modified curriculum. Participant P2.1 was a male aged 14.5 years, and P2.2 was a male aged 14 years. In norm-referenced curriculum-based measurements, Participant P2.1 scored in the 72<sup>nd</sup> percentile for first-grade reading and in the 27<sup>th</sup> percentile for second-grade reading, and he scored in the 54<sup>th</sup> percentile for third-grade math computation. Participant P2.2 scored in the 4<sup>th</sup> percentile for third-grade reading and below the 1<sup>st</sup> percentile for third-grade math computation. Participants were selected for inclusion based on their performance on an adaptation of the relational assessment used in Kirsten and Stewart (2021). All probe and training sessions were administered by the researcher in an otherwise unoccupied room of Participant P2.1's center, and in an unoccupied room at Participant P2.2's house. The researcher sat next to the participant at a desk. A second, independent observer sat approximately 3 feet away from the desk on the other side of the participant with a full view of the participant and the computer screen.

Ethical approval for recruitment of participants was obtained from the director of the clinic, parental consent was obtained for each child who participated, and verbal assent was obtained from each of the participants.

### Experimental Design

As in Experiment 1, a multiple baseline design across participants was used in Experiment 2. Experiment 2 included the same relational preassessment for screening potential participants; a baseline condition in which relating combinatorially entailed relations was tested; a brief pretraining probe condition in which relating directly presented relations (DP analogy) was tested; and a training condition in which relating directly presented relations (DP analogy) was trained and CE analogy generalization probe trials were presented. In order for the second participant to enter the training condition, the first participant had to meet the probe criterion (i.e., scoring at least five out of six [83%] correct on the probe trials).

### Materials and Apparatus

Materials were the same as those used in Experiment 1, and have been described in the Materials section for the previous experiment.

### Procedure

The procedure was identical to that used in Experiment 1 of the present study. The preassessment was administered first in one session and took approximately 20 min to complete. Following preassessment, participants entered the baseline condition of the multiple baseline design, followed by a brief pretraining probe, followed by the intervention condition, which included training and probe sessions. Table 1 (Experiment 1) shows a schematic overview of procedures. An average of eight sets were run per day during the intervention condition; a probe set took on average 3–4 min to complete, and a training set took approximately 1–2 min to complete. Both participants started baseline sessions at the same time, and both participants took approximately 1 week to complete the study once the training condition was implemented (i.e., based on administration of 4–8 probe and training sessions per day).

#### Interobserver Agreement

Procedural fidelity checks and IOA were determined for baseline, probe, and training conditions by a trained research assistant. Procedural fidelity was assessed through the use of a fidelity checklist in which each trial in each condition was scored as either correct or incorrect; correct presentation required adherence to all relevant procedural criteria based on condition and trial type including presentation and use of

the appropriate feedback. Procedural fidelity was assessed for 32% of all trials and was 100%. IOA was calculated on a trial-by-trial basis for each probe and training trial. IOA was assessed for 48% of Participant P2.1's sessions, IOA was 97%; and 21% of Participant P2.2's sessions, IOA was 100%.

## Results and Discussion

### Overview

Following training, both participants successfully showed analogical responding during CE Probe sets, including the original CE Probe Set 1 used during baseline testing, a novel CE Probe Set 2, and the generalization probe, Combinatorially Entailed Probe w/ Distractor (CE+D Probe; see Fig. 9). Both participants scored 100% correct on CE Probe Set 1.

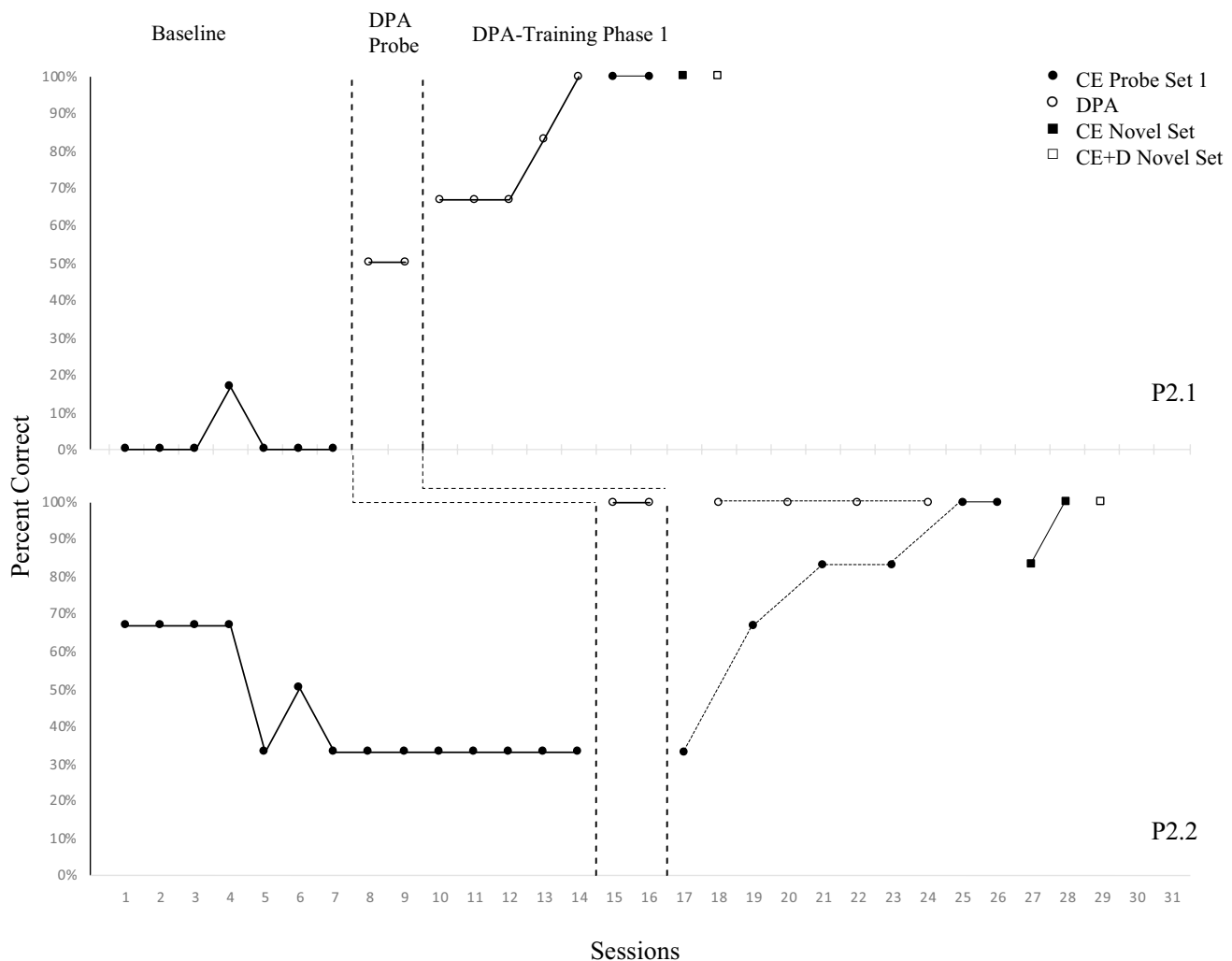
Participant P2.1 scored 100% correct on the novel CE probe, and P2.2 scored 100% correct on his second attempt. Both participants scored 100% correct on the CE+D generalization probe (refer to Table 2 for condition acronyms).

### Preassessment

Both participants were tested for relational responding on the preassessment. Both P2.1 and P2.2 scored 83% and 100% correct, respectively, on the first set of CE trials and did not require a second attempt. Table 4 shows preassessment scores.

### Participant 2.1

Participant P2.1's scores were 0% correct for all but one baseline CE Probe session, and his score was 50% correct



**Fig. 9** Experiment 2: Participant Responding in Analogy Probe and Training Sessions. *Note.* CE Probe Set 1: Combinatorially entailed analogical responses; DP Probe: Directly presented analogy probes; DPA-Training Phase 1: Directly presented analogical training w/ min-

imal feedback; CE+D: Combinatorially entailed analogical responses with distractor; DPA+XF Training Phase 2: Directly presented analogical training w/ increased feedback



on both pretraining DPA-Probe sessions. Participant P2.1 required five DPA-Training sessions before meeting the training passing criteria. He scored 100% on all subsequent probes including two consecutive CE Probe Set 1 sessions, the novel CE Probe, and the CE+Distractor Probe.

### Participant 2.2

Participant P2.2's scores decreased and maintained at low levels of responding during baseline CE Probe sessions, and he scored 100% correct on both pretraining DPA-Probe sessions. However, he scored at baseline level during the first CE Probe after the DPA Probe, thus training was implemented. Participant P2.2 scored 100% correct for all four DPA-Training sessions and he scored 100% correct twice consecutively on CE Probe Set 1 after the fourth training session. Participant P2.2 scored 83% correct on novel CE Probe Set 2 and 100% correct on CE Probe Set 3. Participant P2.2 scored 100% correct on the CE+Distractor Probe.

After direct training in relating directly presented relations, both participants in Experiment 2 showed analogical responding according to RFT's conception of analogy as the derived relating of relations. Neither participant required Phase 2 DPA+XF Training.

The results from the multiple baseline showed that the directly presented (relations) analogy (DPA) training procedure was an effective intervention for training analogy and occasioning generative CE analogical responding as shown by the generalization data in two children with ASD. Both participants passed baseline CE Probe Set 1 as well as a novel CE Probe following DPA analogy training. Furthermore, correct analogical responding generalized to the CE+D Probe.

## General Discussion

The RFT account of analogy as derived relating of relations allows for a functional analysis of analogical responding, which facilitates testing and training of this repertoire. Experiment 1 of the present study aimed to extend previous RFT-based research in analogy in young, typically developing 5-year-old children, and Experiment 2 replicated the procedure with children diagnosed with ASD.

In previous RFT-based research on analogy in young children, Carpentier et al. (2002) found that after testing compound–compound match-to-sample tasks with trained (as opposed to derived) relations, 5-year-old children then successfully passed both equivalence–equivalence (sameness) and nonequivalence–nonequivalence (difference) derived relations tests (i.e., relating combinatorially derived sameness and difference relations). The MTS format used in Carpentier et al. however, posed methodological issues; extensive and laborious pretraining of arbitrary stimuli was required, and the number of potential derived relations based on the initial training network was limited, thus constraining the scope of further testing and generalization as well as of multiple exemplar training if required.

Kirsten et al. (2021) extended Carpentier et al. by using a novel REP type format to test and train analogical relations in 5-year-olds. The REP format required minimal pretraining, and once established, it allowed testing and training of unlimited novel analogies. Kirsten et al. (2021) successfully trained analogy in 5-year-olds using multiple exemplar training in the context of this format. However, unlike Carpentier et al., who tested for relating both sameness and difference relations, Kirsten et al. tested for derived relations between *difference* relations only.

The present study sought to extend the REP methodology used in Kirsten et al. (2021) but with a number of modifications. First, the relational networks included a larger array of relational stimuli than those presented in Kirsten et al., thus permitting tests of relating combinatorially entailed sameness *and* difference relations as had been done in Carpentier et al. (2002). The results from the present study showed that all participants in Experiments 1 and 2 passed the CE analogy probes including sameness and difference relations, as well as the generalization probes.

Second, in the present study, unlike in Kirsten et al. (2021), we did not use multiple exemplar training of deriving relations between derived relations per se in the intervention. Instead, we used an intervention protocol similar in an important respect to that used by Carpentier et al. (2002) in that it involved participants first engaging in the relation of directly presented relations before being tested for the derivation of relations between combinatorially entailed relations. This was similar to Carpentier et al. in that in their study also, participants related directly presented relations

**Table 4** Experiment 2: Preassessment Relational and Analogical Scores (Percent Correct)

Participant	Sort Tokens	CE Sort w/ Tokens Relational	CE Sort w/o Token		CE & Analogy Probe	
			1st Attempt Relational	2nd Attempt Relational	Relational	Analogy
P2.1	100	83	83	N/A	67	0
P2.2	100	100	100	N/A	83	67

*Note.* Inclusion criteria required passing the relational test at 80% or above and failing the analogy test

before being tested for the derivation of relations between derived relations. One key difference in this respect, however, was that in the present experiment most of the participants had to be trained in the relation of directly presented relations rather than being able to engage in this behavior spontaneously as was the case with the participants in Carpentier et al. However, once sufficiently trained in this pattern, all of the participants could subsequently engage in the derivation of relations between derived combinatorially entailed relations without the latter needing to be trained. The fact that the REP format in the present study facilitated the testing and training of a potentially unlimited number of novel analogical tasks permitted an unconstrained quantity of training exemplars as well as of generalization testing, which, as previously noted, contrasted with Carpentier et al. wherein the capacity for doing so was constrained by the MTS methodology.

Experiment 2 of the present study extended the work on analogy to include children diagnosed with ASD. The closest behavioral study of figurative language in children since Carpentier et al. is on metaphorical responding in children diagnosed with ASD (Persicke et al., 2012). The methodology used in Persicke et al., however, did not control for participant history with the stimuli (i.e., potential familiarity with the metaphors themselves or at least stimuli on which they drew) nor did it control for difficulty across metaphor exemplars. The use of arbitrary stimuli in the context of the REP format in the present study obviated the need to control of these variables, and thus allowed us to maintain experimental control while examining analogical responding in children with autism. It is interesting that the children with ASD in Experiment 2 required fewer DPA (directly presented analogy) training trials than the participants in Experiment 1. One possible factor contributing to these results is that the children with ASD may be more familiar with trial-based learning and 1:1 instruction due to their history with applied behavior analytic interventions. Another possible contributing factor is the age difference of nearly 10 years between the typically developing children (approximately 5 years old) and the children with ASD (approximately 14 years old). Regardless of the difference in acquisition, these results indicate that the present format can be used to successfully test and train analogical relations in children with ASD, who characteristically struggle to understand figurative language. Considering that analogy is important not just in itself but also for language and cognition in general, training analogy in children with deficits in language development could result in generativity and creativity in language skills in addition to encouraging intellectual growth. In previous research, Persicke et al. (2012) found that after MET in metaphor, generalization of the ability to comprehend untrained metaphors occurred for all participants with ASD. Furthermore, two of the three participants began to

create their own metaphors during training and posttraining sessions. It is possible that training children with ASD or other developmental delay using procedures such as the present one might similarly result in generalization to the understanding and creation of novel figurative language in a more naturalistic context. This might be a focus of future work.

In both experiments of the present study, only Participant P2.2 in Experiment 2 did not require training on the directly presented analogy tasks as he scored 100% correct in both the DPA (directly presented analogy) probes and training trials. Following success on these trials, additional prompting with this task facilitated correct responding on the relation of combinatorially entailed relations. He required four DPA training sessions, or more accurately DPA prompting sessions, before meeting the probe criteria for relating derived relations. As previously mentioned, one finding of the present study was that three out of four participants required training in relating directly presented relations, which contrasted with the findings in Carpentier et al. (2002). P2.2 of the present study is the only participant who responded correctly to the “relating directly presented relations” tasks without training. This is in contrast to the results in the Carpentier et al. study in which, whereas the children failed initially to show the derivation of relations between derived relation without intervention, they did all spontaneously show derivation of relations between directly trained relations, and of course giving them the latter tasks facilitated their doing the former. It is interesting to speculate as to why three out of the four children in the present study could not spontaneously relate directly presented relations. Perhaps further research could examine whether a difference in the protocols (i.e., MTS vs. REP) produced these contrasting results.

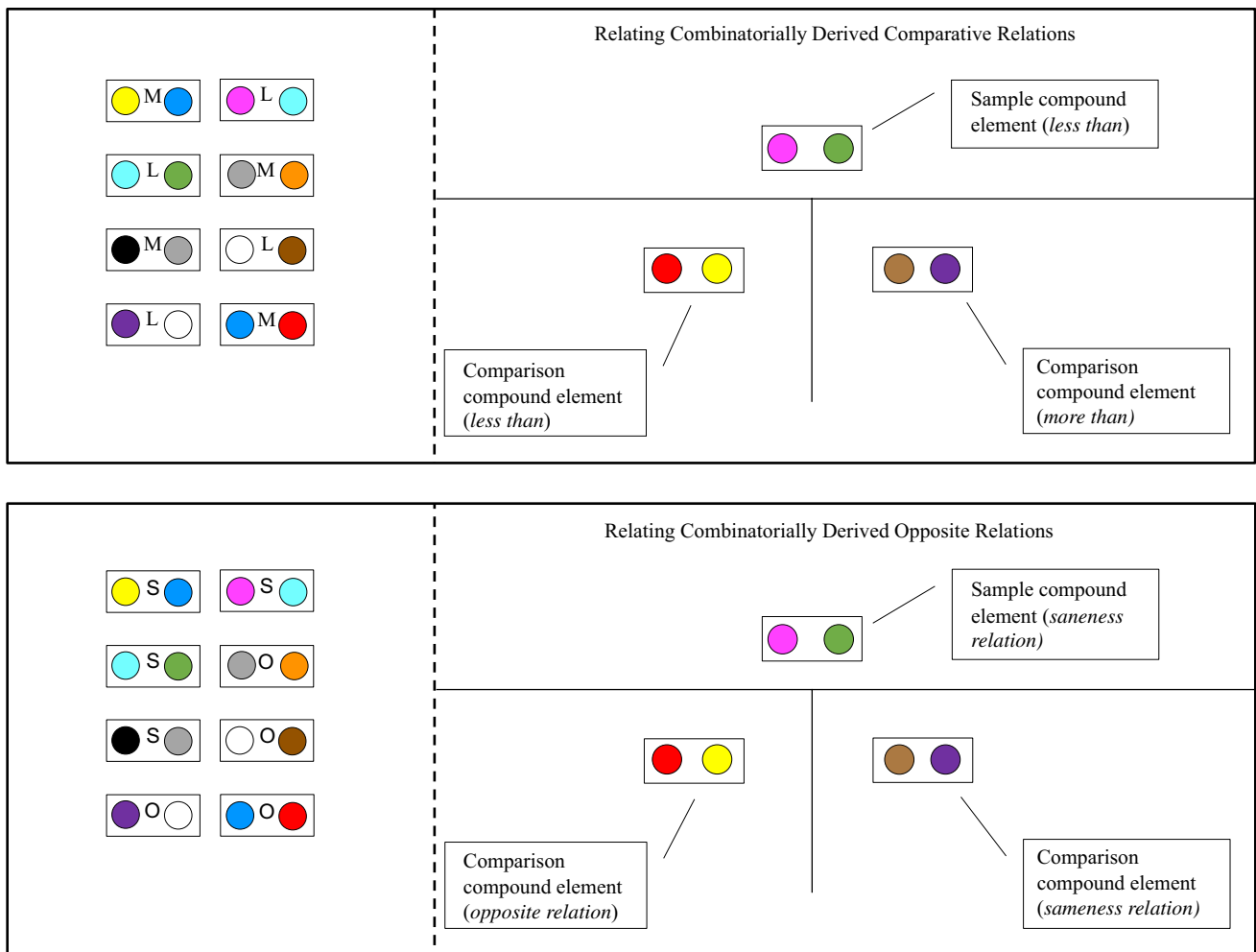
Considering the relevance of analogy to intellectual potential, future researchers could investigate the generalized effects of training analogical responding on socially valid measures such as mainstream analogy tests, academic achievement tests, or standardized tests of cognitive performance. In previous RFT research on intellectual performance and relational responding, Cassidy et al. (2011) and various follow-up studies (e.g., Cassidy et al., 2016; Hayes & Stewart, 2016) used the REP to assess and train derived relational responding and compared scores on pre- and post-training standardized intelligence tests. Participant scores on the intelligence tests increased significantly following the relational training. Future researchers could similarly investigate the effects of training analogy, with multiple different relations within the analogies, on intellectual performance. For example, relational networks could include relations of comparison or opposition and test for relating combinatorial entailed relations as in Fig. 10. In previous research, Lipkens and Hayes (2009) successfully showed analogical

responding across sameness, difference, comparison, and opposite relations in adult participants. A protocol such as that used in the present study might afford the opportunity to efficiently test and train a similar variety of analogies in young children and to subsequently examine the effects of such training on intellectual potential.

A closely related possibility for further research could be to examine the effects of training sameness relations on the emergence of other relations. For example, once participants have been trained in coordinate analogical responding, performance with analogies involving other types of relations (e.g., comparative–comparative) could be tested to see if generalization across relations could occur. Kirsten and Stewart (2021) found that coordinate analogical responding was acquired before comparative, opposite,

temporal, and hierarchical analogical responding. Future research could examine whether relating these other relations might be prompted by training analogy of coordination. An alternative, despite the empirical findings of Kirsten and Stewart, might be to investigate whether under particular circumstances training analogy involving noncoordinate relations might be able to support the emergence of analogy of coordination. MET of analogy might also be tested by examining whether training just one variety of analogy (e.g., coordination) alone facilitates generalization in novel relational varieties of analogy, or whether training additional relational varieties of analogy might be required to promote generalization.

Future researchers investigating the acquisition of analogy might also usefully consider the dimensions along which



**Fig. 10** Examples of Relating Combinatorially Derived Comparative and Opposite Relations. *Note.* Top panel: M = more than; L = less than. In this example, the sample compound depicts a combinatorially entailed less than relation, the left comparison compound depicts a combinatorially entailed less than relation (i.e., the correct response), and the right comparison compound depicts a combinatorially

entailed more than relation. Bottom panel: S = same; O = opposite. In this example, the sample compound depicts a combinatorially entailed sameness relation, the left comparison compound depicts a combinatorially entailed opposite relation, and the right comparison compound depicts a combinatorially entailed sameness relation (i.e., the correct response)

analogical relational responding can vary as described in the multidimensional (and latterly hyperdimensional) multilevel (MDML/HDML) framework of Barnes-Holmes et al. (2017, 2020). The MDML-HDML framework proposes five levels of development of arbitrarily applicable relational responding including mutual entailment, relational framing, relational networking, relating relations (i.e., analogy), and relating relational network and sees these levels as intersecting with four “dimensions” along which relational responding at each of the five levels can vary. The dimensions include relational coherence (the extent to which a given pattern of AARR is in functional agreement within its verbal community), relational complexity (the complexity of a pattern of AARR; e.g., more stimuli mean greater complexity), relational derivation (how “well practiced” a pattern of AARR has become) and flexibility (the extent to which a pattern of AARR may be modified by context). Regarding the focus of the present study, future researchers might refer to the MDML-HDML framework to experimentally analyze how various dimensions intersect with analogy and related levels during acquisition. For example, perhaps children provided with more opportunities to derive relating of relations (i.e., lower levels of derivation) might show improved abilities in the next level up, that is, the relating of relational networks, and a similar point might be made with respect to the training of other dimensions (e.g., relational flexibility).

One potential limitation of the present study was the relatively restricted participant sample. One obvious cause for this was the strict inclusion criteria, which eliminated participants who were unable to fluently derive simple arbitrary relations. Future research along similar lines might consider increasing the age range of the participants involved in order to include a larger sample that might allow better insight into the emergence of analogical responding in young children. Furthermore, the preassessment used in the present study might be considered for further research with regard to testing and training arbitrary relational responding. Further research could examine why some children do not successfully complete the preassessment, and methods for training children how to respond on the REP more effectively could be investigated. Another potential limitation is the age difference between the two populations (i.e., 5-year-old typically developing children vs. 14-year-old children with an ASD diagnosis). However, both participants with ASD scored well below grade level in math and reading norm-referenced curriculum-based measurements (scores now included in the Participants and Setting section) and thus were performing well under the level of typically developing 14-year-old children.

It was also noted that the token procedure used to assess combinatorially entailed relations (CE relations) during the preassessment may warrant further investigation. In Section 3 of the preassessment, participants were required to

sort tokens into sets based on the relational information provided in the relational network (see second panel of Fig. 4). Participants were given 12 monochromatic tokens that matched the colors of the circles in the relational networks, and a sheet of paper divided equally into four parts (see second panel of Fig. 4). One token from each of the four relational networks was placed in one of the four spaces on the paper, and the participants’ task was to sort the remaining eight circles into four sets of three tokens each based on the directly presented, mutually entailed, and combinatorially entailed relations derived from the relational network. This brief and simple exercise obviated the need for more detailed instructions on combinatorial entailment or the function of the compound stimuli required to complete CE trials, including CE analogy trials. Furthermore, informal observations by the researcher suggested that the participants particularly enjoyed this task, including the children who did not participate in the entire study. Future applied-RFT research could examine the efficacy of using manipulable, colored tokens as arbitrary stimuli to assess and train derived multiple relations.

Finally, one additional note might be made regarding the comparison of the REP with MTS. In the foregoing we have touted the advantages of the REP over the MTS. We noted that the MTS format does pose certain methodological issues when assessing or training a relatively complex response pattern such as analogy; for example, extensive and laborious pretraining of arbitrary stimuli is required, and the number of potential derived relations based on the initial training network is limited, thus constraining the scope of further testing and generalization as well as of multiple exemplar training if required. On the other hand, it could be argued that, although more efficient as a protocol once participants are trained on it, the REP does still require initial training in the REP format, and it is possible that for at least some participants (such as those who failed the initial preassessment in the present study for example) that training might pose certain difficulties that perhaps MTS-based training might not. It might also be argued that the use of MTS can allow a more ecologically valid model of analogical reasoning because the required relational responses have to be established in the repertoire of the children before they can be tested. In contrast, with the REP, participants simply have to check the relations on one side of the screen and then respond according to the stimuli presented as analogies on the right hand side of the screen. Thus, although both protocols demonstrate analogical responding, and the REP can be argued to allow much more efficient generation of analogies, it could be argued that the MTS protocol requires that the child has to learn the relational responding (i.e., lower levels of derivation are involved) before being tested for analogical reasoning. Hence, rather than claiming that the REP

is always a better protocol to use in studying analogy (or other complex repertoires) perhaps it might be said that each procedure offers particular advantages depending on the nature of the research and or the particular focus of the study.

The present study adds to the limited behavioral research on analogical responding in young children with and without developmental disabilities, and contributes further evidence that 5-year-old children and children with ASD can be successfully trained in analogical responding. This work further confirms a potential developmental divide in capacity for analogical responding to the extent that the 5-year-olds in the present study were not readily able to show analogy, as well as further highlighting the potential utility of additional training for addressing this deficit. Considering the ubiquity of analogical responding in everyday life, more research regarding its development and training in young children and in children with language delays is merited. Finally, the potential of the REP format used in this study to test and train young children in complex relational responding, such as analogy, is promising and lends itself to further investigation of its experimental and applied utility.

**Availability of Data and Material** The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

**Code Availability** Not applicable

## Declarations

**Conflicts of Interest** The authors declare that they have no conflict of interest.

**Ethical Approval** The study was approved by the NUI Galway Research Ethics Committee. All procedures performed in the studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Consent to Participate** Informed consent was obtained from all individual participants included in the study.

## References

- Alexander, P. A., Willson, V. L., White, C. S., Fuqua, J. D., Clark, G. D., Wilson, A. F., & Kulikowich, J. M. (1989). Development of analogical reasoning in 4- and 5-year-old children. *Cognitive Development*, 4(1), 65–88. [https://doi.org/10.1016/0885-2014\(89\)90005-1](https://doi.org/10.1016/0885-2014(89)90005-1)
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.).
- Barnes, D., Hegarty, N., & Smeets, P. M. (1997). Relating equivalence relations to equivalence relations: A relational framing model of complex human functioning. *Analysis of Verbal Behavior*, 14, 57–83. <https://doi.org/10.1007/BF03392916>
- Barnes-Holmes, D., Hayes, S. C., Dymond, S., & O’Hora, D. (2001). Multiple stimulus relations and the transformation of stimulus functions. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 51–71). Kluwer Academic/Plenum.
- Barnes-Holmes, Y., Barnes-Holmes, D., & Smeets, P. M. (2004a). Establishing relational responding in accordance with opposite as generalized operant behavior in young children. *International Journal of Psychology & Psychological Therapy*, 4, 559–586.
- Barnes-Holmes, Y., Barnes-Holmes, D., Smeets, P. M., Strand, P., & Friman, P. (2004b). Establishing relational responding in accordance with more-than and less-than as a generalized operant behavior in young children. *International Journal of Psychology & Psychological Therapy*, 4, 531–558.
- Barnes-Holmes, D., Regan, D., Barnes-Holmes, Y., Commins, S., Walsh, D., Stewart, I., Smeets, P. M., Whelan, R., & Dymond, S. (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. <https://doi.org/10.1901/jwab.2005.79-04>
- Barnes-Holmes, D., Barnes-Holmes, Y., Luciano, C., & McEntegart, C. (2017). From the IRAP and REC model to a multi-dimensional multi-level framework for analyzing the dynamics of arbitrarily applicable relational responding. *Journal of Contextual Behavioral Science*, 6, 434–445. <https://doi.org/10.1016/j.jcbs.2017.08.001>
- Barnes-Holmes, D., Barnes-Holmes, Y., & McEntegart, C. (2020). Updating RFT (more field than frame) and its implications for process-based therapy. *The Psychological Record*, 70, 605–624. <https://doi.org/10.1007/s40732-019-00372-3>
- Bartha, P. (2019). Analogy and analogical reasoning. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2019 ed.). Stanford University. Retrieved January 16, 2021, from <https://plato.stanford.edu/archives/spr2019/entries/reasoning-analogy/>
- Bassok, M. (2001). Semantic alignments in mathematical word problems. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 401–433). MIT Press.
- Belisle, J., Dixon, M. R., Stanley, C. R., Munoz, B., & Daar, J. H. (2016). Teaching foundational perspective-taking skills to children with autism using the PEAK-T curriculum: Single-reversal “I-You” deictic frames. *Journal of Applied Behavior Analysis*, 49, 965–969. <https://doi.org/10.1002/jaba.324>
- Berens, N. M., & Hayes, S. C. (2007). Arbitrarily applicable comparative relations: Experimental evidence for a relational operant. *Journal of Applied Behavior Analysis*, 40(1), 45–71. <https://doi.org/10.1901/jaba.2007.7-06>
- Bod, R. (2009). From exemplar to grammar: A probabilistic analogy-based model of language learning. *Cognitive Science*, 33(5), 752–793. <https://doi.org/10.1111/j.1551-6709.2009.01031.x>
- Brown, A. L., & Kane, M. J. (1988). Preschool children can learn to transfer: Learning to learn and learning from example. *Cognitive Psychology*, 20(4), 493–523.
- Carpentier, F., Smeets, P. M., & Barnes-Holmes, D. (2002). Matching functionally same relations: Implications for equivalence-equivalence as a model for analogical reasoning. *The Psychological Record*, 52, 351–370. <https://doi.org/10.1007/BF03395435>
- Carpentier, F., Smeets, P. M., & Barnes-Holmes, D. (2003). Equivalence-equivalence as a model of analogy: Further analyses. *The Psychological Record*, 53, 349–371.
- Carpentier, F., Smeets, P. M., Barnes-Holmes, D., & Stewart, I. (2004). Matching derived functionally-same stimulus relations: Equivalence-equivalence and classical analogies. *The Psychological Record*, 54, 255–273. <https://doi.org/10.1007/BF03395473>

- 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. <https://doi.org/10.1007/BF03395755>
- Cassidy, S., Roche, B., Colbert, D., Stewart, I., & Grey, I. M. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learning & Individual Differences*, *47*, 222–235. <https://doi.org/10.1016/j.lindif.2016.03.001>
- Colbert, D., Tyndall, I., Roche, B., & Cassidy, S. (2018). Can SMART training really increase intelligence? A replication study. *Journal of Behavioral Education*, *27*, 509–531. <https://doi.org/10.1007/s10864-018-9302-2>
- Dougher, J. M., Hamilton, D. A., Fink, B. C., & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. *Journal of the Experimental Analysis of Behavior*, *88*(2), 179–197. <https://doi.org/10.1901/jeab.2007.45-05>
- Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more than, and less than. *Journal of the Experimental Analysis of Behavior*, *64*(2), 163–184. <https://doi.org/10.1901/jeab.1995-64-163>
- Dymond, S., & Roche, B. (Eds.). (2013). *Advances in relational frame theory: Research and application*. Context Press/New Harbinger.
- Dymond, S., May, R. J., Munnely, A., & Hoon, A. E. (2010). Evaluating the evidence base for relational frame theory: A citation analysis. *The Behavior Analyst*, *33*, 97–117. <https://doi.org/10.1007/bf03392206>
- Fagot, J., & Maugard, A. (2013). Analogical reasoning in baboons (*papio papio*): Flexible reencoding of the source relation depending on the target relation. *Learning & Behavior*, *41*, 229–237. <https://doi.org/10.3758/s13420-012-0101-7>
- Fryling, M., Rehfeldt, R. A., Tarbox, J., & Hayes, L. J. (Eds.). (2020). *Applied behavior analysis of language and cognition*. Context Press/New Harbinger.
- Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170. [https://doi.org/10.1016/S0364-0213\(83\)80009-3](https://doi.org/10.1016/S0364-0213(83)80009-3)
- Gentner, D., Holyoak, K. J., & Kokinov, B. (Eds.). (2001). *The analogical mind: Perspectives from cognitive science*. MIT Press.
- Gil, E., Luciano, C., Ruiz, F. J., & Valdivia-Salas, S. (2014). A further experimental step in the analysis of hierarchical responding. *International Journal of Psychology & Psychological Therapy*, *14*(2), 137–153.
- Gillan, D. J., Preman, D., & Woodruff, G. (1981). Reasoning in the chimpanzee: I. Analogical reasoning. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*(1), 1–17. <https://doi.org/10.1037/0097-7403.7.1.1>
- Gorham, M., Barnes-Holmes, Y., Barnes-Holmes, D., & Berens, N. (2009). Derived comparative and transitive relations in young children with and without autism. *The Psychological Record*, *59*, 221–246. <https://doi.org/10.1007/BF03395660>
- Goswami, U. (1989). Relational complexity and the development of analogical reasoning. *Cognitive Development*, *4*(3), 251–268. [https://doi.org/10.1016/0885-2014\(89\)90008-7](https://doi.org/10.1016/0885-2014(89)90008-7)
- Goswami, U. (1996). Analogical reasoning and cognitive development. In H. W. Reese (Ed.), *Advances in Child Development and Behavior* (pp. 91–138). Elsevier Science.
- Goswami, U. (2001). Analogical reasoning in children. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 437–470). MIT Press.
- Goswami, U., & Brown, A. L. (1989). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition*, *35*, 69–95. [https://doi.org/10.1016/0010-0277\(90\)90037-K](https://doi.org/10.1016/0010-0277(90)90037-K)
- Goswami, U., & Brown, A. L. (1990). Higher-order structure and relational reasoning: Contrasting analogical and thematic relations. *Cognition*, *36*, 207–226. [https://doi.org/10.1016/0010-0277\(90\)90057-Q](https://doi.org/10.1016/0010-0277(90)90057-Q)
- Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding new verbal capabilities in children with language delays*. Pearson Education.
- 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*, *86*(3), 397–411. <https://doi.org/10.1007/s40732-018-0273-4>
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. Kluwer Academic/Plenum.
- Hofstadter, D. R. (2001). Epilogue: Analogy as the core of cognition. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 499–537). MIT Press.
- Hofstadter, D., & Sander, E. (2013). *Surfaces and essences: Analogy as the fuel and fire of thinking*. Basic Books/Perseus Books Group.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. MIT Press.
- Kalandadze, T., Norbury, C., Naerland, T., & Naess, K. A. B. (2018). Figurative language comprehension in individuals with autism spectrum disorder: A meta-analytic review. *Autism*, *22*(2), 99–117. <https://doi.org/10.1177/1362361316668652>
- Kirsten, E. B., & Stewart, I. (2021). Assessing the development of relational framing in young children. *The Psychological Record*. <https://doi.org/10.1007/s40732-021-00457-y>
- Kirsten, E. B., Stewart, I., & McElwee, J. (2021). Testing and training analogical responding in young children using relational evaluation procedure. *The Psychological Record*. <https://doi.org/10.1007/s40732-021-00468-9>
- Lapiana, W. P. (2004). Merit and diversity: The origins of the Law School Admissions Test. *Saint Louis University Law Journal*, *48*, 955–990.
- Levinson, P. J., & Carpenter, R. L. (1974). An analysis of analogical reasoning in children. *Child Development*, *45*(3), 857–861. <https://doi.org/10.2307/1127862>
- Lipkens, R., & Hayes, S. C. (2009). Producing and recognizing analogical relations. *Journal of the Experimental Analysis of Behavior*, *91*, 105–126. <https://doi.org/10.1901/jeab.2009.91-105>
- Lunzer, E. A. (1965). Problems of formal reasoning in test situations. *Monographs of the Society for Research in Child Development*, *30*, 19–46. <https://doi.org/10.2307/1165774>
- Maenner, M. J., Shaw, K. A., Baio, J., Washington, A., Patrick, M., DiRienzo, M., Christensen, D. L., Wiggins, L. D., Pettygrove, S., Andrews, J. G., Lopez, M., Hudson, A., Baroud, T., Schwenk, Y., White, T., Rosenberg, C. R., Lee, L. C., Harrington, R. A., Huston, M., . . . Dietz, P. M. (2020). Prevalence of autism spectrum disorder among children aged 8 years. *Autism and Developmental Disabilities Monitoring Network*, *69*(SS-4), 1–12. <https://doi.org/10.15585/mmwr.ss6904a>
- McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004). A relational frame account of the development of complex cognitive phenomena: Perspective taking, false belief understanding, and deception. *International Journal of Psychology & Psychological Therapy*, *4*, 303–323.
- Morsanyi, K., & Holyoak, K. J. (2010). Analogical reasoning ability in autistic and typically developing children. *Developmental Science*, *13*(4), 578–587. <https://doi.org/10.1111/j.1467-7687.2009.00915.x>
- Mulhern, T., Stewart, I., & McElwee, J. (2018). Facilitating relational framing of classification in young children. *Journal of Contextual Behavioral Science*, *8*, 55–68. <https://doi.org/10.1016/j.jcbs.2018.04.001>

- O'Hara, D., Pelaez, M., & Barnes-Holmes, D. (2005). Derived relational responding and performance on verbal subtests of the WAIS-III. *The Psychological Record*, *55*, 155–175. <https://doi.org/10.1007/BF03395504>
- Persicke, A., Tarbox, J., Ranick, J., & St. Clair, M. (2012). Establishing metaphorical reasoning in children with autism. *Research in Autism Spectrum Disorders*, *6*, 913–920. <https://doi.org/10.1016/j.rasd.2011.12.007>
- Piaget, J., Montangero, J., & Billeter, J. B. (2001). The formation of analogies (R. L. Campbell, trans.). In J. Piaget, *Studies in reflecting abstraction* (pp. 139–152). Psychology Press. (Original work published 1977)
- Polya, G. (2004). *How to solve it: A new aspect of mathematical method*. Princeton University Press (Original work published 1945).
- Rattermann, M. J., & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a causal-mapping task. *Cognitive Development*, *13*, 453–478.
- Rehfeldt, R. A., & Barnes-Holmes, Y. (Eds.). (2009). *Derived relational responding: Applications for learners with autism and other developmental disabilities*. Context Press/New Harbinger.
- Richland, L. E., & McDonough. (2010). Learning by analogy: Discriminating between potential analogs. *Contemporary Educational Psychology*, *35*, 28–43. <https://doi.org/10.1016/j.cedpsych.2009.09.001>
- Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94*, 249–273. <https://doi.org/10.1016/j.jecp.2006.02.002>
- Ruiz, F. J., & Luciano, C. (2015). Common physical properties among relational networks improve analogy aptness. *Journal of the Experimental Analysis of Behavior*, *103*, 498–510. <https://doi.org/10.1002/jeab.147>
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech & Hearing Research*, *14*(1), 5–13. <https://doi.org/10.1044/jshr.1401.05>
- Smirnova, A., Zorina, Z., Obozova, T., & Wasserman, E. (2015). Crows spontaneously exhibit analogical reasoning. *Current Biology*, *25*, 256–260. <https://doi.org/10.1016/j.cub.2014.11.063>
- Steele, D., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior*, *56*(3), 519–555. <https://doi.org/10.1901/jeab.1991.56-519>
- Sternberg, R. J. (1977). Component processes in analogical reasoning. *Psychological Review*, *84*(4), 353–378.
- Sternberg, R. J., & Nigro, G. (1980). Developmental patterns in the solution of verbal analogies. *Child Development*, *51*(1), 27–38. <https://doi.org/10.2307/1129586>
- Stewart, I. (2016). The fruits of a functional approach for psychological science. *International Journal of Psychology*, *51*(1), 15–27. <https://doi.org/10.1002/ijop.12184>
- Stewart, I., & Roche, B. (2013). Relational frame theory: An overview. In S. Dymond & B. Roche (Eds.), *Advances in relational frame theory: Research and application* (pp. 51–71). Context Press/New Harbinger.
- Stewart, I., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2001). Generating derived relational networks via the abstraction of common physical properties: A possible model of analogical reasoning. *The Psychological Record*, *51*, 381–408.
- Stewart, I., Barnes-Holmes, D., & Roche, B. (2004). A functional-analytic model of analogy using the relational evaluation procedure. *The Psychological Record*, *54*, 531–552. <https://doi.org/10.1007/BF03395491>
- Stewart, I., Barnes-Holmes, D., & Weil, T. (2009). Training analogical reasoning as relational responding. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding: Applications for learners with autism and other developmental disabilities* (pp. 257–279). New Harbinger.
- Stewart, I., McElwee, J., & Ming, S. (2013). Language generativity, response generalization, and derived relational responding. *Analysis of Verbal Behavior*, *29*, 137–155. <https://doi.org/10.1007/BF03393131>
- Stewart, I., McLoughlin, S., Mulhern, T., Ming, S., & Kirsten, E. B. (2020). Assessing and teaching complex relational operants: Analogy and hierarchy. In M. Fryling, R. A. Rehfeldt, J. Tarbox, & L. J. Hayes (Eds.), *Applied behavior analysis of language and cognition* (pp. 198–213). Context Press/New Harbinger.
- Thoma, P., & Daum, I. (2006). Neurocognitive mechanisms of figurative language processing: Evidence from clinical dysfunctions. *Neuroscience & Biobehavioral Reviews*, *30*, 1182–1205. <https://doi.org/10.1016/j.neubiorev.2006.09.001>
- Zettle, R. D., Hayes, S. C., Barnes-Holmes, D., & Biglan, A. (Eds.). (2016). *The Wiley handbook of contextual behavioral science*. Wiley Blackwell.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.