Title	Assessing and training analogical reasoning in young children using relational frame theory
Author(s)	Kirsten, Elle Ben
Publication Date	2021-09-23
Publisher	NUI Galway
Item record	http://hdl.handle.net/10379/17027

Assessing and Training Analogical Reasoning in Young Children Using Relational Frame Theory

Elle Ben Kirsten
HonsBSc Psychology
MSc Psychology

Dissertation submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy in Applied Behaviour Analysis
Supervisor: Dr Ian Stewart
School of Psychology
National University of Ireland, Galway
September 2021

Table of Contents

Declaration Regarding the Work

I, the **Candidate**, certify that the Thesis is all my own work and that I have not obtained a degree in this University or elsewhere on the basis of any of this work.
This thesis is the result of my own investigations, except where otherwise stated.

Signed: ________________________ Date: ___________________

Abstract

Analogical responding is pervasive in everyday language and cognition and is a key component in learning. However, despite this, there is as yet relatively little behavioural research on 1) the age of emergence and 2) assessment and training of this repertoire. The aim of the present thesis was provide additional insight into these issues by extending previous research that used relational frame theory (RFT) to investigate analogical relations in young children. Study 1 aimed to assess the age of acquisition during the typical development of analogical relations in young children between ages three and seven. Studies 2–4 aimed to investigate an RFT-based procedure to test and train analogy in young children; in Experiment 2 of Study 4, the procedure was extended to test and train analogy in children with autism spectrum disorders.

Given the apparent importance of analogy for intellectual development, cognitive-developmental psychologists have examined the emergence of this skill in young children. Early researchers believed that analogical reasoning developed at the age of 12 or later (Levinson & Carpenter, 1974; Lunzer, 1965; Piaget et al., 1977; Sternberg & Nigro, 1980). More recently, however, it has been argued that children as young as four can show analogical reasoning (Goswami & Brown, 1990), with prior knowledge playing a critical role. Research on analogy has been mostly the province of cognitive psychologists, but behaviour analysts have also begun to research analogy during the last two decades. The impetus for this has primarily come from researchers who take an RFT perspective. Carpentier et al. (2002) found that 5-year-olds initially failed to show analogical responding and required additional training before doing so. Study 1 investigated a multi-stage training and testing protocol that allowed for assessing analogical responding in the context of the assessment of participants' relational responding more broadly. Participant analogical relational performance was correlated with their age and intellectual performance as assessed on a standardised test of intellectual functioning.

A second aim of Study 1, a cross-sectional study, was to measure relational responding of various types and at various levels of complexity in order to provide more comprehensive data on the emergence of basic framing patterns in 3- to 7-year-old children. The primary focus was on analogy, or the relating of relations, as one particularly important pattern of relational responding. Relational frame theory (RFT) views the operant acquisition of various patterns of relational framing (frames) as key to linguistic and cognitive development, and it has explored the emergence of a range of psychological phenomena (e.g., analogy,

perspective-taking) in these terms. Despite the growing evidence that relational framing 1) underlies human cognition and language and 2) is operant behaviour, there is little research on the normative development of relations in young children. One potentially important advance for RFT research is to obtain more detailed information on the normative development of relational framing in childhood. Study 1 examined a range of frames, including coordination, comparison, opposition, temporality, and hierarchy at four different levels of complexity, of which two levels looked specifically at analogical responding (nonarbitrary relating, *nonarbitrary relating of relations*, arbitrarily applicable relating, and *arbitrarily applicable relating of relations*). The relational evaluation procedure (REP)-based training and testing format utilised in the current study was employed in the context of the multi-stage protocol relational assessment that allowed testing for a range of different types of relations.

Study 2 assessed and trained analogical responding in young, typically developing children. Three 5-year-old children were assessed and trained in relating relations using an RFT-based REP protocol in a combination multiple baseline design across participants and a multiple probe design across behaviours. The study included a relational pre-assessment to screen potential participants; a baseline condition in which analogy was tested; and a training condition in which analogical responding was trained and generalisation probe trials (including three different probes; CE, DMC, and D-Cue Probes) were presented. After training in relating combinatorially entailed relations, all three participants showed analogical responding according to RFT's conception of analogy as the derived relating of relations.

Study 3 was a replication of Study 2; however, the correction procedure was modified to provide more training opportunities for incorrect responding. In Study 2, the probe trials were presented whether the participant emitted a correct or incorrect response during the training correction procedure. In Study 3, the training procedure required participants to respond correctly before the probe was re-presented. Three 5-year-old children were assessed and trained in relating relations using the RFT-based REP protocol in a combination multiple baseline design across participants and a multiple probe design across behaviours. As in Study 2, following multiple exemplar training, correct responding increased to criterion levels for all three children, and both generalisation and maintenance were observed.

In Study 4, the testing and training procedure was modified to include a larger array of stimuli, and directly trained and mutually entailed relations within relations were assessed in addition to combinatorially entailed relations as in Studies 2 and 3. In Experiment 1 of Study 4, two 5-year-old typically developing children were assessed and trained in relating relations in a multiple baseline design. 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 generalisation probe, CE+D Probe. Experiment 2 was a replication of Experiment 1; however, Experiment 2 sought to investigate analogical responding in children diagnosed with autism spectrum disorders. Two children with ASD were assessed and trained in relating relations. 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 generalisation probe, CE+D Probe. These results suggest that this format can be used to successfully train children with ASD to respond to analogical relations as defined by RFT.

The present thesis offers further insight into the testing and training of anlogical relations in young children. The analogy assessment data provide further evidence that analogical relations develop around age five, and the training data show that analogy can be successfully trained when the repertoire is weak or missing in five-year-old children.

The data from the present studies suggest the experimental and applied potential of the REP format. The REP format permitted multiply controlled studies in which we could target analogy testing and training directly while maintaining experimental control. Furthermore, it afforded us with quick and effective stimulus control, allowing us to implement multiple baseline designs to examine the efficacy of multiple exemplar training to establish the core repertoire. Future work could extend the REP format to test and train analogical relations beyond coordination and distinction. A closely related possibility for further research could be to examine the effects of training sameness relations on the emergence of other relations.

The present thesis contributes to the extant behavioural research on relational language assessment and training. The resulting data suggest that regardless of the level of complexity of the derivation required (i.e., whether directly presented, mutually or combinatorially entailed), relations between relations should be considered analogy. Thus, derived relations between nonderived relations are also analogies, albeit simpler than derived relations between derived relations. The present data also constitute an important addition to the literature on analogical relations beyond behaviour analysis. Although cognitive researchers have an extensive literature on analogy, they have not yet presented a functional analytic model analogy that might lend itself to training this repertoire.

Considering the relevance of analogy to intellectual potential, future researchers could investigate the generalised effects of training analogical responding on socially valid measures such as mainstream analogy tests, academic achievement tests, or standardised tests of cognitive performance. Future work could examine if training children with ASD or other developmental delays using procedures such as the present one might result in generalisation to the understanding and creation of novel figurative language in a more naturalistic context. Considering the potential for improving language and cognition, further research into relational assessment and training, in general, is undoubtedly warranted.

To my baby brother, Ben.
Your absence is the presence that follows me everywhere.
You will remain in my heart forever.
May 2, 1996 – August 14, 2016

Acknowledgements

My sincere gratitude to:

Dr Ian Stewart. Thank you for always being so patient and kind during this Ph.D., it threw some curve balls at me. Your supervision and advisement have been invaluable - so much work has come out of our conversations about analogy, RFT, and language testing and training. All of the children I worked with during my Ph.D. are indebted to you for some of their more impressive language gains! I am so grateful that I had this opportunity to work with you. Thank you!

Mom. Sidekicks for life! I am pretty sure you are the only person who unintentionally understands the concept of arbitrarily applicable relational responding. Thank you for sticking with me through everything.

Dad. Finishing this project would have been much harder/impossible without your continuous support. I wish you could see the impact you have had on so many lives because of all your generosity; there are many children speaking with meaning and listening with understanding because of you. We are so very grateful to you.

Johann. Yīnyáng. Thanks for the many chats, the constant support, and the laughs. And, of course, the countless cool/fun/fancy stuff you keep sending my way. You make life easier and sweeter. Thank you! Be water, my friend.

Sanmarie. The Secret Language of Jellyfish Dancers. Ópera Sobre Basura. Niemand anders sal dit regtig kry nie, least of all because of our accents. Tú sabes. Thanks for literally sharing a laugh ;)

Ruth, Laura, and Alexandra. My soulmates, you are so precious to me. I am beyond grateful to each of you for sharing your love, wisdom, kindness, and magic. Your friendships are my lifeline.

Ruth – You + Lake Mahopac, Debz, Grass Days, and The Rock. From Blue to Rainbow, I love you.

Laura – Vic’s and Shuka, Domaine Eden, and the Brooklyn Botanical Gardens… You bring so much colour and flavour to my world.

Kaitlin. Maybe in our next lives, we won’t be the old-men muppets (Karma, are you listening?!)! So grateful for all your support and laughs through all of it.

Adam, Maddy, Jasmine, and Liz. Thank you for the IOA, the feedback, the piloting, the patience, and the laughs!

Drs Nick and Kim Berens. Thank you for the playground – none of this would have been possible without the space you created for me to learn, practice, and explore. Thank you!

Lauren G-B. Your support made this project so much easier than it would have been. I cannot ever thank you enough. SSDS—Toda raba!

Dr Sarah Cassidy. Your research inspired me so much I ended up in Ireland! I’m so grateful for the opportunities I have had to learn about RFT and ACT from you. Thank you!

Dr Heidi Eilers. Thank you for so generously sharing your ACT expertise with me and for teaching me to how apply ACT in my practice! Talk about transformation of stimulus function!

Dr Laura Crosby. Another soulmate! I look forward to our weekly chats more than you know, thank you so much for recognizing our connection. I love you!

Dr David Gast. Sometimes I felt crazy—thank you for sharing your time, humour, and research-design-knowledge so generously.

Dr Teresa Mulhern. Thank you for holding my hand since day one!

Let the beauty of what you love be what you do. – Rumi

To all of the kids and parents I have had the pleasure of working with throughout all my academic and clinical adventures—thank you for letting me be a part of your lives; you are the reasons I love this work!

“Because of the dog’s joyfulness, our own is increased. It is no small gift.” – M. Oliver

To my precious angel babies, Furrow and Tweety - life would be very boring without your velvety snoots, warm cuddles, and silly adventures. And to all my pups who, at some point, slept (snored) next to me while I was typing: Smalley, Charlie Bear, Jackie, Lucy, Selby, Lady Lovebug, and Rocky, big love for always, my furry lovebugs.

Jennifer. 4142 Seasons – I’m here for all of it. I love you for always.

Publications and Conference Presentations

Publications

Stewart, I., McLoughlin, S., Mulhern, T., Ming, S., & Kirsten, E. B. (2020). Assessing and teaching complex relational operants: Analogy and hierarchy. In R. Rehfeldt, J. Tarbox, M. Fryling, & L. Hayes (Eds.), *Applied Behavior Analysis of Language and Cognition*. New Harbinger.

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 Frame Theory. *The Psychological Record*. https://doi.org/10.1007/s40732-021-00468-9

Kirsten, E. B., Stewart, I., & McElwee, J. (2022). Testing and Training Analogical Relational Responding in Children with and Without Autism. *The Psychological Record*. https://doi.org/10.1007/s40732-021-00493-8

Conference Presentations and Posters

Kirsten, E., & Stewart, I. (2017, April). *Developing and Pilot Testing a Protocol to Measure Analogical Relational Responding in Young Children* [Conference symposium]. Experimental Analysis of Behaviour Group, London, UK.

Kirsten, E. B., Berens, K. N., & Berens, N. M. (2017, May 25-29). *Relational frame theory in practice: Producing generative language in applied settings* [Conference poster]. 43th Annual Convention of the Association for Behavior Analysis, Denver, CO.

Kirsten, E. B., & Stewart, I. (2018). *Developing and Testing an Assessment of Analogical Responding in Young Children* [Conference symposium]. Psychological Society of Ireland Division of Behaviour Analysis, Galway, Ireland.

Stewart, I., Bast, D., & Kirsten, E. B. (2018). *Assessing and Training Young Children Using Relational Frame Theory* [Conference workshop]. Psychological Society of Ireland Division of Behaviour Analysis, Galway, Ireland.

Ming, S., Stewart, I., McElwee, J., Bast, D., & Kirsten, E. B. (2018, July 24-29). *A Functional Contextualist Approach to Early Language Training: Using Relational Frame Theory to Promote Linguistic Generativity* [Conference workshop]. Association of Contextual Behavioral Science 2018 World Conference, Montreal, Canada.

Kirsten, E. B., & Stewart, I. (2018, July 24-29). *Advances in Relational Frame Theory Research of Applied Relevance* [Conference symposium]. Association of Contextual Behavioral Science 2018 World Conference, Montreal, Canada.

Kirsten, E. B., & Stewart, I. (2018). *The Analogical Relations Assessment: Assessing Relational Frames and Analogical Responding in Young Children* [Conference paper]. Chester CBS 2018 Research Colloquium, Chester, UK.

Kirsten, E. B., Theil, M., Hockman, A., LaBarbera, K., & Berens, K. (2018, November). *Changing Channels: Building Flexibility in Articulation, Spelling, and Relational Language* [Conference paper]. 31st Annual International Precision Teaching 2018 Conference, Seattle, WA.

Kirsten, E. B., & Stewart, I. (2019, April 15-17). *Methodological Advances in Testing and Training Analogical Responding in Young Children* [Conference symposium]. Experimental Analysis of Behaviour Group, London, UK.

Kirsten, E. B., & Stewart, I. (2019). *Testing and training analogy in young children using a novel RFT-based methodology* [Conference symposium]. Psychological Society of Ireland Division of Behaviour Analysis 2019, Galway, Ireland.

Kirsten, E. B., & Stewart, I. (2019, June 25-30). *Developing Training Protocols to Test and Train Analogical Responding in Young Children* [Conference symposium]. Association of Contextual Behavioral Science 2019 World Conference, Dublin, Ireland.

Kirsten, E., & LaBarbera, K. (2019, November). *Changing Channels Part II: Spelling and Analogical Relations* [Conference symposium]. 32nd Annual International Precision Teaching Conference 2019, St. Petersburg, FL

Kirsten, E., & Stewart, I. (2020, May 21-25). *Assessing Relational Responding in Young Children Using a Novel Relational Frame Theory-Based Relational Evaluation Procedure-Based Format* [Conference symposium]. 46th Annual Convention of the Association of Applied Behavior Analysis 2020, Washington D.C/Online.

Kirsten, E., & Stewart, I. (2020, May 21-25). *Training Analogical Responding in Young Children Across Several Multiple Baseline Design Studies* [Conference symposium]. 46th Annual Convention of the Association of Applied Behavior Analysis 2020, Washington D.C/Online.

Kirsten, E., & Stewart, I. (2020, July 16-19). *Modern Considerations for Relational Frame Theory and Contextual Behavioural Science; Conceptual and Empirical Advances* [Conference symposium]. Association of Contextual Behavioral Science 2020 World Conference, Online.

Kirsten, E., & Stewart, I. (2020). *The Normative Development of Relational Framing* [Conference presentation]. Psychological Society of Ireland Division of Behaviour Analysis 2020, Ireland.

Kirsten, E., & Stewart, I. (2020). *Testing and Training Analogical Relations in Young Children* [Conference presentation]. Psychological Society of Ireland Division of Behaviour Analysis 2020, Ireland.

Kirsten, E., Cassidy, S., Ryan, A., & Laly, L. (2021, June 24–27). *ACT in Action: Examining cutting edge modalities of delivering ACT to meet higher demand for services* [Conference symposium]. Association of Contextual Behavioral Science 2021 World Conference, Online.

List of Figures

List of Tables

List of Appendices

List of Abbreviations

A	After
AA	Arbitrary analogical relations
AADRR	Arbitrarily applicable derived relational responding
AARR	Arbitrarily applicable relational responding
ABA	Applied Behaviour Analysis
ABLLS-R	Assessment of Basic Language and Learning Skills–Revised
APA	American Psychiatric Association
AR	Arbitrary relations
ASD	Autism spectrum disorder
B	Before
C	Contains
CE	Combinatorial entailment
CE+D	Combinatorially entailed plus a distractor
D	Different
D-Cue	Difference cue probe
DMC	Directly presented/mutually entailed/combinatorially entailed probe
DP	Directly presented
DPA	Directly presented analogy
DPA+XF	Directly presented analogy plus extra feedback
FR	Fluid Reasoning
FSIQ	Full-Scale IQ
HDML	Hyper-level multi-dimensional
I	Inside
IOA	Interobserver agreement
IQ	Intelligence quotient
KN	Knowledge
L	Less
LSAT	Law School Admissions Test
M	More
MBD	Multiple baseline design
ME	Mutual entailment
MET	Multiple exemplar training

MTS	Match-to-sample
NAA	Nonarbitrary analogical relations
NAARR	Nonarbitrarily applicable relational responding
NAR	Nonarbitrary relations
NVIQ	Nonverbal IQ
O	Opposite
PEAK	Promoting the Emergence of Advanced Knowledge
PEAK-E-PA	Promoting the Emergence of Advanced Knowledge Equivalence Pre-Assessment
QR	Quantitative Reasoning
RA	Relational assessment
REP	Relational Evaluation Procedure
RFT	Relational Frame Theory
S	Same
S-	Stage
SS-	Substages
SB5	Stanford-Binet Intelligence Scales (5th Edition) for Early Childhood
T	Total
V	Verbal
VB-MAPP	Verbal Behavior Milestones and Placement Programs
VIQ	Verbal IQ
Viz	Visual
Voc	Vocal
VS	Visual-Spatial Processing
WJ III GIA	Woodcock-Johnson III Tests of Cognitive Abilities – General Intellectual Ability – Standard
WM	Working Memory
WPPSI-R	Wechsler Preschool and Primary Scale of Intelligence – Revised

Chapter 1: Overview

Analogy is central to learning in both children and adults and is considered a critical skill for further knowledge acquisition (Gentner, 1988; Gentner & Rattermann, 1991; Goswami, 1996; Hofstadter & Sander, 2012; Holyoak & Thagard, 1995; Morsanyi & Holyoak, 2010; Polya, 1945/2004; Sternberg, 1977; Stewart et al., 2004, 2013). For example, consider the analogy often cited to explain subatomic particles, ‘an atom is like the solar system’. The relation of the electrons to the nucleus is brought into an equivalence relation with the relation of the planets to the sun, ‘electrons orbit the nucleus much like planets orbit a star’. This comparison shows the basic process involved in analogy, whereby an individual’s familiarity with a known domain (e.g., the solar system) can be used to teach them about important aspects belonging to a second unknown domain (e.g., atomic structure). Furthermore, analogy is not solely confined to education or science but is evident in situations as diverse as learning figurative speech (e.g., a mother telling her belligerent child that they are walking on ‘thin ice’) or understanding new technologies. For example, in 1981, Apple compared their new computer to a bicycle. In this analogy, the computer is to the mind as the bicycle is to a person, meaning our computer amplifies the mind as the bicycle amplifies a person (who is looking to travel with speed and efficiency; Hey et al., 2008).

Analogical reasoning lies at the core of intelligence and creativity (Bod, 2009; Gentner, 1983; Green et al., 2012; Hofstadter, 2001; Hofstadter & Sander, 2013; Holyoak & Thagard, 1995; Oppenheimer, 1956; Sternberg, 1977). It is a key component of higher-order language and cognition, including scientific and mathematical skills (e.g., Matos & Passos, 2010; Polya, 1945/2004; Richland & Simms, 2015; Sternberg, 1977), as well as problem-solving more generally (e.g., Brown, 1989), and it is commonly used as a metric of intellectual potential (e.g., Sternberg, 1977) and as a measure to predict academic success; for example, in the Law School Admissions Test (LSAT; Lapiana, 2004).

Research on analogy has been mostly the province of cognitive psychologists (Alexander, 1989; Goswami, 1989; Goswami & Brown, 1990; Levinson & Carpenter, 1974; Lunzer, 1965; Piaget et al., 1977/2001; Sternberg & Nigro, 1980). However, during the last two decades, behavioural psychologists have also begun to research analogy. Specifically, relational frame theory (RFT), a contemporary, functional account of language and cognition, has explicitly recognised the theoretical importance of analogy (see, e.g., Stewart, Barnes-Holmes & Weil, 2009) and has provided a functional model of analogical responding as the ability to derive relations between relations. This functional analytic definition of analogical responding has allowed RFT researchers to experimentally investigate analogy. One strand of research of particular interest for the present purpose focused on analogical responding in young children (e.g., Carpentier et al., 2002, 2003).

Despite the apparent importance of analogical reasoning within higher cognition, a functional analytic assessment and training protocol targeting this repertoire has yet to be established. Relational frame theory's account of language as arbitrarily applicable relational responding provides an empirically sound framework for investigating analogical responding (Dymond & Roche, 2013; Fryling et al., 2020; Hayes et al., 2001; Rehfeldt & Barnes-Holmes, 2009; Zettle et al., 2016). According to RFT, humans respond to stimuli in terms of others via the control of contextual cues signalling how the stimuli are related, independent of their formal or topographical properties. These contextual cues, or relational frames, provide a functional analytic scaffold for investigating language, including analogical responding. The main goals of this thesis were to investigate: 1) the development of relational framing, including analogical relations, in young children, and 2) the assessment and training of analogical relations in young children from an RFT perspective.

Before presenting the series of studies in the current thesis, the relevant cognitive and behavioural background research will be discussed in Chapter 2.

Chapter 3 presents Study 1, a correlational study that sought to measure relational responding of various types, and at multiple levels of complexity, in young children across a range of age groups. A second aim of the study was to focus in particular on analogy, or the relating of relations, as one particularly important pattern of relational responding. Study 1 examined a range of relational frames including coordination, comparison, opposition, temporality, and hierarchy at different levels of complexity (nonarbitrary relating, nonarbitrary relating of relations, arbitrarily applicable relating and arbitrarily applicable relating of relations) in young children ranging in age from 3 to 7 years.

To date, very little research on the assessment and training of analogical relations in young children exists in the behavioural literature. In Chapters 4 and 5 (Studies 2 and 3) typically developing children were successfully assessed and trained in analogical responding using an RFT-based protocol, the relational evaluation procedure (REP), in a multiple baseline design.

Chapter 6 presents Study 4, in which the training protocol used in Studies 2 and 3 is modified. In Experiment 1 of Study 4, typically developing children are assessed and trained in analogical responding. In Experiment 2, the procedure is successfully replicated with children diagnosed with autism spectrum disorder.

A general discussion of our findings and implications for future research are presented in Chapter 7.

Chapter 2: Introduction

Figurative language, including analogy, is an integral component of all languages' basic linguistic structure (Dancygier & Sweetser, 2014). Analogies are ubiquitous in daily language and cognition and lie at the core of intelligence and creativity (Bod, 2009; Gentner, 1983; Green et al., 2012; Hofstadter, 2001; Hofstadter & Sander, 2013; Holyoak et al., 2001; Holyoak & Thagard, 1995; Oppenheimer, 1956; Sternberg, 1977a, 1977b). Analogy is central to learning in both children and adults (Alexander et al., 1989; Gentner & Rattermann, 1991; Goswami, 1996; Hofstadter & Sander, 2012; Holyoak & Thagard, 1995; Morsanyi & Holyoak, 2010; Polya, 1945/2004; Sternberg, 1977a; Richland & Simms, 2015; Stewart et al., 2004, 2013). It is a key component of higher-order language and cognition, including scientific and mathematical skills (e.g., Polya, 1945/2004; Sternberg, 1977a) as well as problem-solving more generally (e.g., Brown, 1989), and it is commonly used as a metric of intellectual potential (e.g., Sternberg, 1977a), and as a measure to predict academic success; for example, in the Law School Admissions Test (LSAT) (Lapiana, 2004). Reasoning by analogy is considered a critical skill for further knowledge acquisition and language generativity (Gentner, 1988; Gentner & Rattermann, 1991; Goswami, 1996; Matos & Passos, 2010). Furthermore, through analogy, we can construct convincing legal arguments based on set precedents, or convey complicated emotions through poetry and prose (Gick & Holyoak, 1980; Spellman & Schauer, 2005). For example, a literal explanation of a subject's personal qualities such as kindness, charm, and beauty would not evoke the same emotive behaviour as Shakespeare does when he writes: 'Shall I compare thee to a summer's day?' (Sonnet 18; Dancygier & Sweetser, 2014). 'Emotional experiences are notoriously difficult or impossible to convey by literal language; but by connecting the relational pattern of a novel experience with that of a familiar, emotion-laden one, analogy provides a way of recreating a complex pattern of feelings' (Holyoak et al., 2001, p. 5). And indeed, cultural records provide prolific examples of analogy in literature, religion, and philosophy.

What Is Analogy?

The term analogy is borrowed from the Greek analogia, a term used by Greek mathematicians to denote a similarity in proportional relationships (Hesse, 1965; Stewart & Barnes-Holmes, 2001). Aristotle first studied proportional analogies as a form of logical reasoning, as demonstrated by his classic syllogism, 'All men are mortal; Socrates is a man; ergo, Socrates is mortal' (Hofstadter & Sander, 2013, p. 17). Aristotle's classical four-term analogical structure A:B::C:D depicts this equality of proportion. Mathematical in their precision, these proportional analogies are often included in intelligence tests but arguably do not include the more enlightening analogies inspiring scientific discoveries (e.g., penicillin) or facilitating complex concept explanation (e.g., atomic structure via a planet and sun analogue), influence creative design, poetry, humour, empathy, political debate, and so forth (Holyoak, 2005; Holyoak et al., 2001). Instead, analogies of attribution were another form of analogy identified by the Greeks in which the similarity of function was inferred by was that of inferring similarity of function in which the two analogues are linked by a common property attributed in some way to each term (Stewart et al., 2001, p. 75). In Metaphysics, Aristotle summarised both types of analogy as A is in B or to B, so that C is in D or to D (Theta IX, Ch. 6, 1048b). The essence of analogical thinking is the transfer of knowledge from one domain, A : B, to another, C : D, via reference to aspects of one body of information and aspects of another (Gick & Holyoak, 1983). Thus, essentially, terms are analogous if they have properties in common or when there is a similarity in relation (Hesse, 1965).

Within the field of psychology itself, there is little agreement on reasoning by analogy between subdisciplines. This should come as no surprise considering the fundamental lack of unity in psychology wherein each subdiscipline offers its own theories and data analyses with limited generalisation of principles between theoretical approaches (Chiesa, 1994). Regarding analogy, the cognitive sciences have long been interested in analogy, particularly its development in young children. More recently, behaviour psychology, specifically relational frame theory, has provided a functional analytic definition of analogy, and the development of analogy in young children has been investigated (Barnes et al., 1997; Carpentier et al., 2002, 2003, 2004).

Cognitive Science

Analogical reasoning is the relating of two situations based on sharing a common pattern of relationships among their component elements (e.g., Holyoak, 2005). The situations may belong to the same domain (i.e., within-domain analogy or literal analogies, focusing on the use of old problems to solve new ones; i.e., involves items that belong to the same or at least very close conceptual domains) or may pertain to two unrelated domains (i.e., cross-domain analogy or metaphorical analogies; two fundamentally different or remote conceptual domains that share a similar explanatory structure. For example, the analogy between the atom and the solar system is based on the similarity in the structure of the two systems (e.g., sun-nucleus, electrons-planets) (Vosniadou, 1989; Vosniadou & Ortony, 1989). Both types of analogies usually transfer information from the known domain to the novel domain. The distinction between within domain and between domain analogies represents a continuum of comparisons involving items that are clear examples of the same concept to items along with different remote domains; analogical reasoning can be employed between items that belong anywhere in the continuum from literal similarity to non-literal similarity (Vosniadou, 1989).

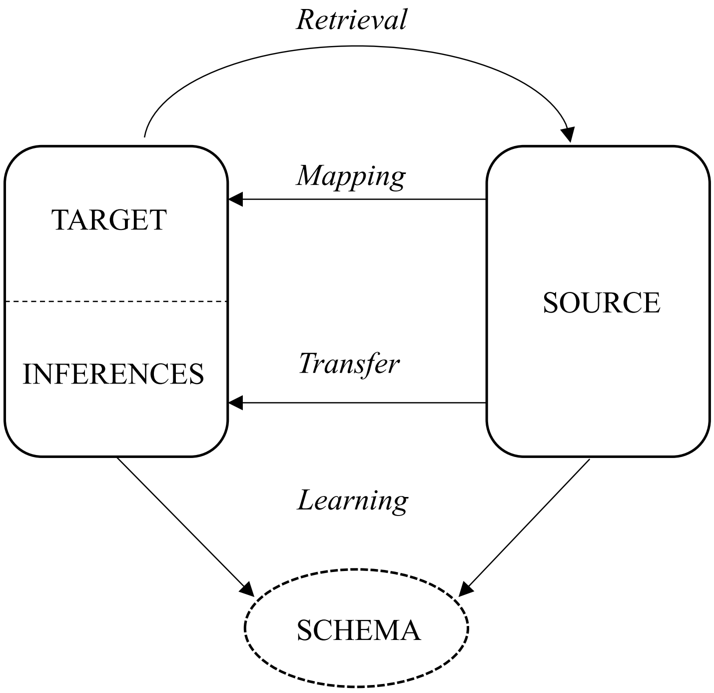
Hesse (1966) described analogy as involving two distinct relationships: the 'vertical' relationship between the two parts of each analogue and the 'horizontal' mapping relationship between the two analogues. That is, an analogy is a hierarchy of relations in which there is a higher-order relation of equivalence between two lower-order relations (Sternberg, 1977b). Analogical reasoning is the ability to identify and compare the relations between two things; it is the process of understanding a novel situation in terms of a familiar one. The familiar situation provides a model for making inferences about the unfamiliar situation, and through analogical reasoning, the novel situation evolves into another type of example of the known familiar situation (Holyoak et al., 2001). This type of generativity is a critical distinction of analogical reasoning and is regarded as one of the most sophisticated aspects of human cognition (Holyoak & Thagard, 1995; Sternberg, 1977b; Stewart, Barnes-Holmes, Hayes, & Lipkens, 2001).

Analogy became of interest to mainstream cognitive researchers when they realised that human reasoning does not always operate based on content-free general inference rules but rather is often related to particular bodies of knowledge and the context in which it occurs (Vosniadou & Ortony, 1989). Within the cognitive sciences, analogy is defined as an 'inductive mechanism based on structured comparisons of mental representations' (Holyoak, 2005, p. 234). Two situations are analogous if they share a common pattern of relationships among their constituent elements despite the elements being different across the two situations (Holyoak, 2005). Identifying a common pattern requires a comparison of the two situations. In most cases, one analogue, the source or base, is more familiar than the other analogue, the target, in that prior experience or knowledge about functional relations within the source analogue are known, for example, understanding that certain properties of the source have causal, explanatory, or logical connections to other properties (Hesse, 1996). The initial asymmetry in knowledge supports the process of analogical transfer, an inductive process, in which the source is used to generate inferences about the target (Holyoak, 2005).

There is general agreement that an analogy requires a relational structure normally applied in one domain to also function in another domain (Gentner, 1983); whole systems of connected relations are matched from a known domain that already exists in memory (i.e., the source, the base, or the vehicle analogue) to another, unknown domain (the target) (Holyoak et al., 2001; Gentner et al., 2001; Vosniadou & Ortony, 1989). The general consensus among cognitive scientists is that analogical thinking can be deconstructed into several basic component processes, including: 1) one or more relevant analogues stored in memory must be accessed; 2) a familiar analogue must be mapped to the target analogue, and the corresponding parts of each analogue must be aligned; 3) analogical inferences are made from the mappings, allowing new knowledge to fill gaps in understanding; 4) the inferences are evaluated and possibly adapted to fit the unique requirements of the target; 5) new categories and schemas may be generated as a result of the analogical reasoning (Gentner & Holyoak, 1997; Holyoak et al., 2001; Vosniadou & Ortony, 1989). See Figure 2.1 for an illustration of the processes. All cognitive theories of analogy include and emphasise one or more of these basic component processes. According to cognitive scientists, these component processes involve dynamic interactions of many interrelated systems and thus, as is characteristic of the cognitive science approach to dealing with complex information-processing problems, computational simulations of analogical reasoning have been developed to test analogical theories (Vosniadou & Ortony, 1983).

Figure 2.1

Cognitive Psychology's Basic Component Process in Analogical Reasoning



Gentner (1983) argued that the key similarities in the domains in an analogy lie in the relations within each domain and that analogical similarities often rely on higher-order relations, or relations between relations, such as causal relations (Holyoak et al., 2001). Parenthetically, in RFT, this might be called relating relations within relations. Her structure-mapping theory aims to capture the essential elements that constitute analogy and the operations required in processing analogy (Gentner, 1989). In structure mapping, the process relies on structural/relational commonalities as opposed to specific prior content and is not influenced by the system's problem-solving goals, except to the extent that these goals affect the current representation of the domain.

In her structure-mapping theory, Gentner set forth the view that analogy requires identifying a structural alignment, referred to as the mapping, between domains. The mapping between two representational structures is characterised by structural parallelism, the consistent, one-to-one correspondences between mapped elements, and systematicity, the implicit preference for profound, interconnected relations governed by higher-order relations, such as causal, mathematical, or functional relations (Gentner & Holyoak, 1997; Holyoak et al., 2001). The mapping process takes as input two structured representations, the base (sometimes called source) and target and computes one or more mappings. Each mapping consists of a set of correspondences, each linking a particular item (entity or statement) in the base with a particular item (entity or statement) in the target (Gentner & Forbus, 2011). Structure-mapping theory was instantiated in computer simulations of analogical mapping and inference—the structure-mapping engine and analogical retrieval (the MAC/FAC programmes) (Holyoak et al., 2001; Markman & Gentner, 1993).

Holyoak proposed that analogy offered a more profound approach to problem-solving (Gentner & Holyoak, 1997; Gick & Holyoak, 1983; Holyoak et al., 2001). Holyoak investigated the role of pragmatics in analogy in complex cognitive tasks—how context influences the interpretation of analogy (Holyoak et al., 2001). Gick and Holyoak (1983) found that analogy facilitated the formation of new relational categories by abstracting the relational correspondences into a schema. Holyoak and Thagard (1989) proposed a multiconstraint approach to analogy in which interactions between similarity, structural parallelism, and pragmatic factors or purpose between the source and the target produced an interpretation. The multiconstraint theory assumes that analogical reasoning is directed by several general constraints that cooperatively encourage coherence in analogical thinking (Holyoak & Thagard, 1997). First, the analogy is guided by the similarity of the elements involved. This similarity of concepts at any level of abstraction may contribute to analogical thinking. Second, the analogy is guided by the requirement to identify consistent structural parallels between the source and target domains (Gentner, 1983). Third, the constraint of purpose implies that analogical reasoning is guided by context and goals; for example, what is the purpose of the analogy (Holyoak & Thagard, 1997). These three kinds of constraints do not operate independently; they vary in convergence and opposition and are in constant dynamic interaction pressing toward an internal coherent compromise.

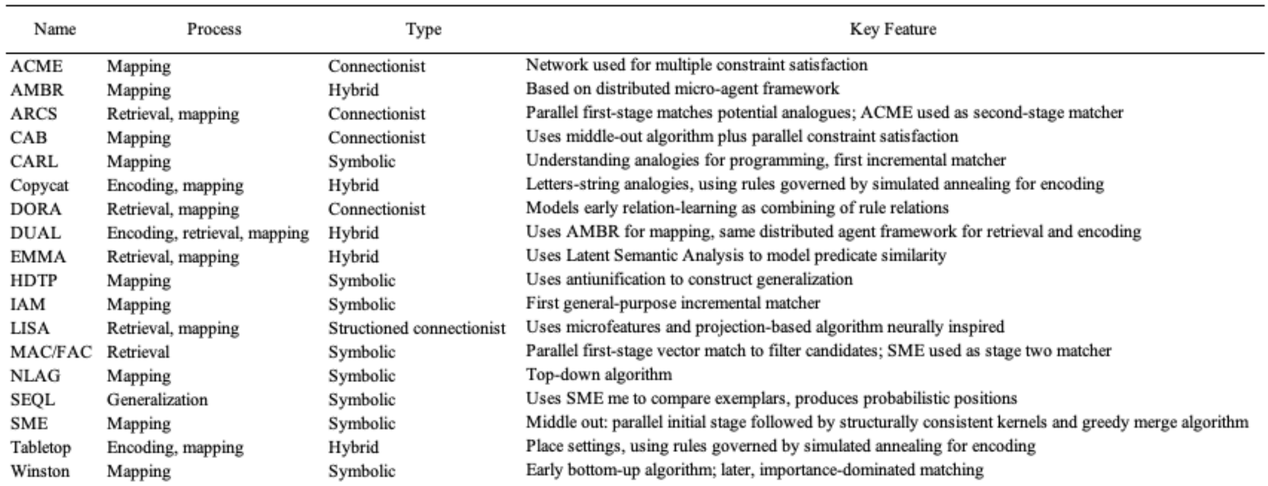
They also developed computational simulations for the theoretical model of analogical mapping and inference (Analogical Constraint Mapping Engine, ACME) based on algorithms for satisfying simultaneous constraints of supporting and competing hypotheses regarding which elements to map to find the overall mapping that best fits the interacting constraints. Hummel and Holyoak (1997) extended the multiconstraint theory and developed a new computer simulation, Learning and Inference with Schemas and Analogies (LISA), which included representations and processing assumptions more consistent with the operation of human memory.

Artificial intelligence scientists and cognitive psychologists have created a number of computational models of analogy to provide theoretical illustrations of how humans compare representations, retrieve potential analogues from memory, and learn from the results

(Gentner & Forbus, 2011). Computational models of analogy include subsets of the basic components processes offered by cognitive science. However, computational models differ in their focus; some capture the range of analogical phenomena at the cognitive level, and others suggest how analogical processes might be implemented in neural systems. Some recent work has focused on modelling interactions between analogy and other processes and modelling analogy as part of larger cognitive systems (Gentner & Forbus, 2011). For example, connectionist models, such as LISA, are more flexible with regard to semantic content and are neurologically plausible (Stewart et al., 2004). For a brief summary of computational models, refer to Table 2.1.

Table 2.1

Computational Models of Analogy and Their Key Characteristics



Name	Process	Type	Key Feature
ACME	Mapping	Connectionist	Network used for multiple constraint satisfaction
AMBR	Mapping	Hybrid	Based on distributed micro-agent framework
ARCS	Retrieval, mapping	Connectionist	Parallel first-stage matches potential analogues; ACME used as second-stage matcher
CAB	Mapping	Connectionist	Uses middle-out algorithm plus parallel constraint satisfaction
CARL	Mapping	Symbolic	Understanding analogies for programming, first incremental matcher
Copycat	Encoding, mapping	Hybrid	Letters-string analogies, using rules governed by simulated annealing for encoding
DORA	Retrieval, mapping	Connectionist	Models early relation-learning as combining of rule relations
DUAL	Encoding, retrieval, mapping	Hybrid	Uses AMBR for mapping, same distributed agent framework for retrieval and encoding
EMMA	Retrieval, mapping	Hybrid	Uses Latent Semantic Analysis to model predicate similarity
HDTP	Mapping	Symbolic	Uses antiunification to construct generalization
IAM	Mapping	Symbolic	First general-purpose incremental matcher
LISA	Retrieval, mapping	Structioned connectionist	Uses microfeatures and projection-based algorithm neurally inspired
MAC/FAC	Retrieval	Symbolic	Parallel first-stage vector match to filter candidates; SME used as stage two matcher
NLAG	Mapping	Symbolic	Top-down algorithm
SEQL	Generalization	Symbolic	Uses SME me to compare exemplars, produces probabilistic positions
SME	Mapping	Symbolic	Middle out: parallel initial stage followed by structurally consistent kernels and greedy merge algorithm
Tabletop	Encoding, mapping	Hybrid	Place settings, using rules governed by simulated annealing for encoding
Winston	Mapping	Symbolic	Early bottom-up algorithm; later, importance-dominated matching

Note. Adapted from ‘Computational Models of Analogy,’ by D. Gentner and K. D. Forbus, 2011, Cognitive Science, 2, p. 268 https://doi.org/10.1002/wcs.105

From a functional behaviour analytic perspective, a fundamental weakness shared by both the connectionist and the representational models is that they do not have functional definitions for the analogical reasoning but rely on information-processing concepts such as ‘mapping’ and ‘knowledge transfer’. From a behavioural perspective, the use of psychological terms, such as matching and mapping, does not explain the core relational performances. Furthermore, despite connectionist models having advantages over the representational models, they are arguably more interesting as models of neurological rather than psychological functioning (Stewart et al., 2004). Regarding the computational models, these terms are better defined, but their similarity to the human psychological events can only be assumed (Stewart et al., 2001).

Behavioural Science

Research on analogy has been mostly the province of cognitive psychologists; however, during the last two decades, behavioural psychologists have also begun to research analogy. For example, relational frame theory (RFT) literature has explicitly recognised the theoretical importance of analogy (see, e.g., Stewart, Barnes-Holmes & Weil, 2009). From a behavioural perspective, the cognitive models of analogies do not describe the relational processes required for analogical reasoning and instead use lay terms such as matching, mapping, and transfer to define a procedure or outcome (Stewart et al., 2001). However, until relatively recently, behaviour analysis had not provided a viable alternative to the cognitivist approach. Skinner (1957) provided an interpretative account of analogy as a form of ‘metaphorical extension,’ a subtype of the ‘extended tact’ (Stewart et al., 2004). A tact is defined as ‘a verbal operant in which a response of a given form is evoked (or at least strengthened) by a particular object or event or property of an object or event’ (Skinner, 1957, pp. 81-82). The tact is a verbal response that makes contact with nonverbal stimuli in the environment for which the speaker receives generalised reinforcement such as verbal praise. For example, a child sees a dog and says, ‘Dog!’, and consequently receives praise from the listener, ‘That’s right, it’s a dog!’ The discriminative stimulus is the dog and the child makes verbal contact with the environment by saying ‘dog’. The listener (the parent) reinforces the child (the speaker) with generalised reinforcement (praise). The tact allows the speaker to infer something about his environment which has nothing to do with himself. The extended tact is a more complex verbal behaviour that occurs when a response is evoked by a novel stimulus that resembles a stimulus previously present when a response was reinforced. Metaphorical verbal behaviour is a subtype of extended tact that occurs ‘because of the control exercised by properties of the stimulus which, though present at reinforcement, do not enter into the contingency respected by the verbal community’ (p. 92). The following is an example of the Skinnerian interpretation of metaphorical extension that appears in ‘Verbal Behavior’ (1957):

> When for the first time a speaker calls someone a mouse, we account for the response by noting certain properties—smallness, timidity, silent movement and so on—which are common to the kind of situation in which the response is characteristically reinforced and to the particular situation in which the response is now emitted. Since these are not the properties used by zoologists or by the lay community as the usual basis for reinforcing a response we call the extension metaphorical (p. 93).

Skinner conceptualises analogy as the abstraction, via the extended tact, of a common physical property, from two different types of environmental events. This conceptualisation initiated the study of analogy within the behavioural sciences, but additional accounts of the behavioural processes underlying this phenomenon are required. For example, how does ‘A is to B as C is to D’ develop from a (presumably) simpler repertoire of formal property abstraction? This is one important question left unanswered by the Skinnerian analysis. In addition, empirical analyses of analogy are required in order to provide a more complete functional analytic treatment of analogical responding.

RFT is a contextual behavioural account of human language and cognition (Hughes & Barnes-Holmes, 2016), which sees arbitrarily applicable derived relational responding (AADRR) or relational framing as the key operant underlying these repertoires (Hayes, Barnes-Holmes, & Roche, 2001; Stewart, 2016; Stewart & Roche, 2013). AADRR is the learned skill of responding to one event in terms of another, based on contextual cues that specify the relation rather than on the formal or physical properties of the stimuli being related. As initially demonstrated by Sidman (1971), derived sameness or equivalence is the earliest and best-researched example of AADRR. In this study, Sidman showed the untrained emergence of a pattern of equivalence relations among picture, text, and spoken word stimuli in an individual with substantial linguistic deficits. Since that seminal research, others have provided evidence that derived equivalence is a learned and trainable pattern of responding (e.g., Luciano, Gomez, & Rodriguez, 2007). In addition, RFT research has also provided evidence of various other (non-equivalence) patterns of AADRR or relational frames including, for example, difference (e.g., Steele & Hayes, 1991), opposition (Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Dymond et al., 2008), comparison (Berens & Hayes, 2007; Dymond & Barnes, 1995), hierarchy (Gil et al., 2014; Ming et al., 2018), analogy (Barnes et al., 1997), temporality (O’Hora et al., 2005), and deixis (McHugh et al., 2004) amongst others.

Relational frame theory (RFT) is a contextual behavioural account of human language and cognition (Hayes et al., 2001; Stewart, 2016; Stewart & Roche, 2013) that views arbitrarily applicable derived relational responding (AADRR), or relational framing, as the key operant underlying these repertoires. Many species can be trained to engage in nonarbitrarily applicable relational responding (NAARR), which involves relating stimuli based on their physical properties (e.g., selecting a stimulus physically similar, different, smaller, or larger than another). However, humans alone can learn AADRR, which involves relating stimuli based on contextual cues that specify the relation rather than on the formal or physical properties of those stimuli.

According to RFT, all forms of framing are characterised by three properties, namely, mutual and combinatorial entailment and transformation of functions. Mutual entailment is the property of bi-directionality of stimulus relations whereby if A is related to B, then B will be related to A (e.g., if A is larger than B, then B is smaller than A). Combinatorial entailment involves the combination of previously acquired relations to allow derivation of novel relations (e.g., if A is larger than B and B is larger than C, then A is larger than C, and C is smaller than A). Transformation of stimulus functions is the property whereby the psychological functions of a stimulus in a derived relation change depending on the nature of the relation and the functions of related stimuli. For example, if a child has already learnt that A has monetary value and then derives that a novel coin C is worth more than A, they will choose C over A although C is novel. Alternatively, if an arbitrary stimulus A has acquired an aversive function (e.g., by being paired with shock), then, despite not being directly paired with shock itself, C might become more aversive than A based on the derived 'larger than' relation. Indeed, previous research has shown such an effect (e.g., Dougher et al., 2007). RFT regards arbitrarily applicable derived relational responding as generalised operant behaviour acquired via multiple exemplar training (MET) provided through typical exposure to a child's socio-verbal environment.

Empirical analyses of relational framing as a generalised operant show that it can be trained with targeted multiple exemplar training if it is weak or absent (Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Barnes-Holmes, Barnes-Holmes, Smeets, Strand, & Friman, 2004; Berens & Hayes, 2007; Dymond et al., 2010; McHugh et al., 2004; Steele & Hayes, 1991; Stewart, 2016). For example, Barnes-Holmes, Barnes-Holmes, Smeets, Strand, and Friman (2004) successfully trained comparative AADRR in 4- to 6-year-old children. The researchers used two or three identically sized paper coins of assorted colours to test and train patterns of relational responding in accordance with more than and less than relations. During the experiment, they explained to the children how the coins compared to each other in terms of their value and then asked them to pick the coin that would buy the most candy. Using MET, they successfully established more-less relations as generalised operant behaviour by reinforcing correct answers (see also Berens & Hayes, 2007; Gorham et al., 2009).

More recent studies have shown that training relational framing also affects general cognitive ability. For example, Cassidy et al. (2011) investigated the effects of an automated multiple-exemplar relational training programme for frames of coordination, opposition, and comparison. Participants included typically developing 10- to 12-year-old children, several of whom were experiencing educational difficulties in school. Cassidy et al. found significant increases in IQ scores for all participants who received relational training. These data clearly indicate that cognitive ability is strongly related to one's repertoire of derived relational responding. Several replications of Cassidy et al. (2011) have shown similar increases in IQ scores following relational training (e.g., Cassidy et al., 2016; Colbert et al., 2018; Hayes & Stewart, 2016).

The studies just reviewed constitute part of a growing body of evidence suggesting that relational framing is the principal behaviour characterising human language and cognition, such that to understand how humans acquire relational framing is to understand how they acquire language (Dymond & Roche, 2013; Fryling et al., 2020; Zettle et al., 2016). One other source of evidence is research that has used laboratory-controlled patterns of relational framing to model and investigate analogy as one particular subtype of language. 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 *peach is to pineapple as goat is to horse*. In this case, peach and pineapple participate in an equivalence relation in the context of fruit; and goat and horse participate in an equivalence relation in the context of animal, and thus, because these are both equivalence relations, we can derive a relation of equivalence between the relations themselves.

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 equivalence relations, called ‘equivalence-equivalence’ responding. For example, consider the analogy apple is to orange as dog is to sheep. In this case, apple and orange participate in an equivalence relation in the context of fruit and dog and sheep participate in an equivalence relation in the context of animal, and thus, because these are both equivalence relations, we can derive a relation of equivalence between the relations themselves. In order to empirically model this phenomenon, Barnes et al. first trained and tested four three-member equivalence relations in adults and 9-year-old children. They used a matching-to-sample (MTS) procedure to train conditional discriminations amongst three-letter nonsense syllables (coded using alphanumeric designations) as follows: A1→B1, A1→C1, A2→B2, A2→C2, A3→B3, A3→C3, A4→B4, A4→C4 and then tested for the derivation of the following four untrained relations: B1↔C1, B2↔C2, B3↔C3, B4↔C4. After participants passed these equivalence tests, they were then tested for the derivation of equivalence relations between equivalence (and non-equivalence) relations themselves (i.e., equivalence-equivalence responding). This involved using compound stimuli comprising either two nonsense syllables that were equivalent or two that were non-equivalent. Participants were required to choose an equivalent pair in the presence of an equivalent pair (i.e., equivalence-equivalence) and a non-equivalent pair in the presence of a non-equivalent pair (i.e., non-equivalence-non-equivalence). For example, given B3C3 and B3C4 as comparisons, if the sample was B1C1, then they had to choose B3C3, whereas if B1C2 was the sample, they had to choose B3C4. All participants related equivalence relations to other equivalence relations and non-equivalence relations to other non-equivalence relations, and thus this constituted a basic model of analogical reasoning.

Stewart et al. (2001) extended Barnes et al. by demonstrating equivalence-equivalence responding based on the abstraction of common formal properties. Using an MTS procedure, participants were taught to choose a specific nonsense syllable in the presence of each of four blue and four red geometric shapes and then to choose a further nonsense syllable in the presence of each of the first eight. During testing, participants demonstrated equivalence responding based on the abstraction of colour by consistently matching nonsense syllables related to same-coloured shapes to each other. Participants also showed equivalence-equivalence responding in which equivalence relations from the previous part of the experiment were related to other equivalence relations, and non-equivalence relations were related to other non-equivalence relations.

Stewart et al. (2002) demonstrated that relating derived relations could allow for the discrimination of common physical similarities between relations and that this subsequently led to a transformation of functions. Tasks were designed such that equivalence-equivalence responding might allow participants to discriminate a physical similarity between the relations involved. Some participants (colour subjects) received only equivalence-equivalence tasks in which they may discriminate a colour relation, whereas others (shape subjects) were given tasks in which they might discriminate a shape relation. A control group received both types of tasks. Stewart et al. trained and tested adults for the formation of four three-member equivalence relations: A1-B1-C1, A2-B2-C2, A3-B3-C3, and A4-B4-C4. The B and C stimuli were three-letter nonsense syllables, and the A stimulus was a coloured shape. Participants successfully tested for equivalence-equivalence responding (e.g., matching B1/C1 to B2/C2 rather than B3/C4). In a subsequent test for the discrimination of formal similarity, colour subjects matched according to colour, shape subjects matched according to shape. Participants who first sorted a series of wooden blocks according to colour instead of shape were exposed to an analogical protocol in which all trials required the discrimination of common shapes across the equivalence-equivalence network. Participants modified their block sorting so that they sorted according to shape. Stewart et al. suggested this modelled the experience of ‘insight’ via analogy, whereby analogical responding facilitates a new, more effective response to the environment. The control group showed no consistent matching pattern.

Stewart et al. (2004) used the relational evaluation procedure (REP), to test and train analogical relations. The experiment involved 9 stages in which 5 participants completed a complex series of REP training and testing protocols. At the end of the training, all participants demonstrated 24 completely novel instances of responding

Barnes-Holmes et al. (2005) tested analogical reasoning based on the relating of derived sameness and derived difference relations. In Experiment 1, Barnes-Holmes et al. recorded reaction time measures of similar-similar (e.g., ‘apple is to orange as dog is to cat’) versus different-different (e.g., ‘he is to his brother as chalk is to cheese’) derived relational responding, in both speed-contingent and speed-noncontingent conditions. In Experiment 2, Barnes-Holmes et al. examined the event-related potentials (ERPs) associated with these two patterns of responding. Both experiments found similar-similar responding to be significantly faster than different-different responding. The behavioural and neurophysiological data suggest that similar-similar responding is simpler and functionally distinct from different-different analogical responding.

Lipkens and Hayes (2009) examined topography-based responses and additional relations, including nonsymmetrical ones. In Experiment 1, participants successfully recognised 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. Lipkens and Hayes found the procedures occasioned relational responses consistent with an RFT account, including productive responding based on analogies.

Ruiz and Luciano (2011) examined cross-domain analogy as relating relations among separate relational networks by correlating participant performance with a standard measure of analogical reasoning. In two experiments, adult participants were first administered general intelligence and analogical reasoning tests. Next, participants completed computerised conditional discrimination training designed to establish two relational networks, each consisting of two 3-member equivalence classes. Testing included a two-part analogical test in which participants had to relate combinatorial relations of coordination and distinction between the two relational networks. In both experiments, 65% of participants passed the analogical test on the first attempt, and results from the training procedure were strongly correlated with the standard measure of analogical reasoning.

Miguel et al. (2015) investigated whether tact training would establish analogies measured by equivalence-equivalence relations. In Experiment 1, college students were trained to tact ‘same’ or ‘different’ in the presence of AB and BC compounds based on component class membership (e.g., A1B1 as ‘same’ and A1B2 as ‘different’), and were then tested on emergent tacts (BA, CB, AC, CA) and equivalence-equivalence relations. Only one of six participants passed all tests without remedial training. In Experiment 2, six college students were trained to tact only compounds belonging to the same class as ‘same’. Three of six participants passed all tests without remedial training. In Experiment 3, six college students were trained to tact stimuli belonging to the same class with a common name prior to exposure to relational tact training. All participants passed tests without remedial training. In Experiment 4, eight college students were trained to tact stimuli belonging to the same class with a common name. Six participants passed without remedial training, while two, who did not tact the relation of the compounds, did not. Results from these studies suggest that simple discrimination of individual components and their relation in the form of tacts is related to equivalence performance.

Ruiz and Luciano (2015) examined analogical ‘aptness’. Twenty participants were trained to respond to the structure of analogical tests, after which they were trained on two separate relational networks, each consisting of three equivalence classes (Network: 1 F1-G1-H1, F2-G2-H2, F3-G3-H3; Network 2: M1-N1-O1, M2-N2-O2, M3-N3-O3). The node stimuli always appeared with colour spots on their backgrounds (F1 and M1: yellow; F2 and M2: red; F3 and M3: blue). During testing, participants were to select the more accurate response from two options: relating combinatorial relations of coordination with the same colour in the node stimuli (e.g., relating G1H1 to N1O1) versus relating combinatorial relations with different colours in the node stimuli (e.g., relating G1H1 to N2O2). The colours of the node stimuli did not appear on the test. Eighteen participants selected the analogies with common colour properties as the more correct ones.

Meyer et al. (2019) evaluated the effects of listener training on the emergence of analogical reasoning, as measured via equivalence-equivalence, and explored the role of verbal behaviour when solving analogy-type tasks. Eighteen college students were trained to select component stimuli from 2 classes, labelled ‘vek’ and ‘zog.’ Tacts and relational responding in the presence of baseline (AB and BC), symmetry (BA and CB), and transitivity (AC and CA) compounds were evaluated. In Experiment 1, 5 out of 6 participants passed analogy tests, but none of them engaged in the relational tacts ‘same’ and ‘different’ during tact tests, possibly due to lack of instructional control. A change in instructions during Experiment 2 produced relational tacts in 4 of 6 participants, and 5 participants passed analogy tests. In Experiment 3, we implemented a talk-aloud procedure to determine if the participants were emitting relational tacts during analogy tests. All 6 participants tacted stimuli relationally and engaged in problem-solving statements to solve analogy tests. Results from these studies suggest that listener and speaker behaviour in the form of relational tacts and other problem-solving statements influenced the participants’ equivalence-equivalence performance.

For the present thesis, the most relevant extension of Barnes et al. (1997) is research assessing derived relations between relations in young children. In the wake of Barnes et al. (1997), one stream of research used their equivalence-equivalence model to investigate the emergence of analogy in young children (Carpentier et al., 2002, 2003). Carpentier et al. (2002) used the equivalence-equivalence paradigm to investigate analogy in a range of age groups, including adults, 9-year-old, and 5-year-old children. As in the original Barnes et al. study, they found that adults and 9-year-old participants readily showed equivalence-equivalence responding. In contrast, the 5-year-old children, while readily passing equivalence testing, initially failed to show equivalence-equivalence responding and required additional training before doing so. More specifically, they required training and testing with compound-compound-matching tasks with trained relations (e.g., A1B1-A3B3 and A1B2-A1B3) before they could successfully pass the derived compound relations (BC-BC) test. Carpentier et al. (2003) extended this work by assessing whether this additional training could also facilitate the 5-year-old children’s ability to pass equivalence-equivalence tests before receiving the prior equivalence tests. This was something that Barnes et al. had shown that adults and 9-year-old children could do, and this was replicated by Carpentier et al. (2003). However, despite providing considerable additional training, only two of 18 of the 5-year-old participants were successful in this task. The Carpentier et al. (2002, 2003) studies thus provided additional insight into the development of equivalence-equivalence responding as a functional analytic model of analogy. By providing a precise, functional analytic model of this behaviour, this work has arguably shed additional light on this phenomenon beyond that provided by mainstream, cognitive psychological work by not alone confirming a developmental divide in the analogical ability at a particular age but also suggesting how additional training might remediate in this respect.

Except for Stewart et al. (2004), 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 derived relations or 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 behavioural, 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 (i) establish relations of sameness and difference amongst arbitrary nonsense syllables and (ii) 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 utilised 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 programme (see also Cassidy et al., 2016; Hayes & Stewart, 2016). The automated programme 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.

In Study 1 of the present thesis, we 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 coloured 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. For a testing analogy, compound stimuli (i.e., one sample compound and two comparison compounds) composed of coloured circles in either same or difference relations were presented below a relational network (see Figure 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 non-readers, to report on and evaluate multiple exemplars of arbitrarily applicable relational networks defined by specifically selected contextual cues.

Studies 2 and 3 of the present thesis sought to extend Carpentier et al. (2002) and further examine the acquisition of analogical responding operationalised as derived relations between relations in 5-year-old children. It extended the earlier work first by attempting to directly train the skill of deriving relations between relations using a controlled multiple baseline design to demonstrate experimental control. In addition, in the present study, we used an alternative, more efficient format to assess and train the derived relational pattern than the match-to-sample (MTS) procedure used by Carpentier et al. Although often used in studies of derived relations, MTS procedures require extensive baseline training before any testing or training can begin.

In Studies 2 and 3 of the present thesis, we adapted the arbitrarily applicable relational stages of the REP-based assessment designed in Study 1 to test and train analogical relations in 5-year-old children . We 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 Study 3 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 Figure 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) trained and tested for both combinatorially derived sameness and difference relations. Study 3, therefore, suggested that in future research in this domain, the array of stimuli in the relational network should be increased to allow for both combinatorially derived sameness and difference relations.

In Experiment 1 of Study 4, we sought to extend Studies 2 and 3 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.

In Experiment 2 of Study 4, we replicated Experiment 1 but extended it to children with autism spectrum disorders (ASD). Characterised by impairments in social interaction and social communication (American Psychiatric Association, 2013), ASD currently affects one 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 behavioural 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 generalised 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 Study 4 of the present thesis, 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.

In summary, analogy is an important type of emergent intellectual ability. RFT has examined and modelled analogy in young children and provided an important contribution to understanding the acquisition of this ability. RFT, thus, has offered a substantive model of analogy and has identified analogy as an important emergent repertoire. The present thesis aimed to further examine analogical responding from an RFT perspective.

Chapter 3: Study 1. Investigating the Acquisition of Analogy and Relational Framing More Generally in Young Children: A Correlational Study

Portions of this chapter have been published:

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.

Introduction

Analogy, the coordination of functionally similar sets of relations, is pervasive in human language and critically important in various key domains of human activity, including science, technology, and education; furthermore, it is frequently used as a metric of intelligent behaviour (e.g., the Miller Analogies Test, the Graduate Record Examination; Gentner, 1983; Morsanyi & Holyoak, 2010; Sternberg, 1977, Stewart, Barnes-Holmes, & Roche, 2004).

Given the apparent importance of analogy for intellectual development, cognitive-developmental psychologists have examined the emergence of this skill in 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 (Levinson & Carpenter, 1974; Lunzer, 1965; Piaget, Montangero, & Billeter, 1977; Sternberg & Nigro, 1980). More recently, however, it has been argued that children as young as four can show analogical reasoning (Goswami & Brown, 1990), with prior knowledge playing a critical role. For example, Goswami and Brown (1989) examined the effect of children's previous experiences and found that children as young as three could complete analogical tasks when they had relevant knowledge about the domains involved (see also Alexander et al., 1989; Goswami, 1989).

While research on analogy has been mostly the province of cognitive psychologists, during the last two decades, behaviour analysts have also begun to research analogy. The impetus for this has primarily come from researchers who take a Relational Frame Theory (RFT) perspective. Working within an RFT framework, Barnes, Hegarty, and Smeets (1997) provided the first functional analytic definition of analogy as the derivation of a sameness or equivalence relation between equivalence relations, called 'equivalence-equivalence' responding. In order to provide an empirical model of this phenomenon, Barnes et al. first trained and tested four three-member equivalence relations in adults and nine-year-old children. After participants passed equivalence tests, they were then tested for the derivation of equivalence relations between equivalence (and non-equivalence) relations themselves (i.e., equivalence-equivalence responding). All participants related equivalence relations to other equivalence relations and non-equivalence relations to other non-equivalence relations, and thus this constituted a basic model of analogical reasoning.

Since Barnes-Holmes et al., this equivalence-equivalence model has been extended to investigate analogy in terms of several important properties, including reliance on the abstraction of common formal properties (Ruiz & Luciano, 2015; Stewart, Barnes-Holmes, Roche, & Smeets, 2001), neurophysiological activity across levels of analogical difficulty (Barnes-Holmes et al., 2005), and the emergence of analogy in young children (Carpentier et al., 2002, 2003). Given the focus of the present paper, we will concentrate on the latter work with children.

Carpentier et al. (2002) used the equivalence-equivalence paradigm to investigate analogy in a range of age groups, including adults, 9-year-olds, and 5-year-olds. As in the original Barnes et al. study, they found that adults and 9-year-olds readily showed equivalence-equivalence responding. In contrast, the 5-year-olds, while readily passing equivalence testing, initially failed to show equivalence-equivalence responding and required additional training before doing so. More specifically, they required training and testing with compound-compound matching tasks with trained relations before they could successfully pass the derived compound relations test. Carpentier et al. (2003) extended this work by assessing whether this additional training could also facilitate the 5-year-olds' ability to pass equivalence-equivalence tests before receiving the prior equivalence tests. However, despite providing considerable additional training, only 2 of 18 of the 5-year-old participants were successful in this task. The Carpentier et al. (2002, 2003) studies thus provided additional insight into the development of equivalence-equivalence responding as a functional analytic model of analogy. By providing a precise, functional-analytic model of this behaviour, this work has arguably shed additional light on this phenomenon beyond that provided by mainstream, cognitive psychological work by not alone confirming a developmental divide in the analogical ability at a particular age but also suggesting how additional training might remediate in this respect.

As the foregoing suggests, the RFT concept of the derivation of relations provides a useful vehicle for studying analogical reasoning. Carpentier et al. employed this approach to shed additional light on the acquisition of analogy in young children. The purpose of the present study was to extend Carpentier et al. in a number of key ways as well as to extend on previous research on the testing of derived relational responding more generally by employing a multi-stage protocol (the Analogical Relational Assessment; ARA).

First, Carpentier et al. examined analogical reasoning in several age groups but only tested two groups of children, namely 5- and 9-year olds. The present study used the derived relations between relations approach to study analogical responding in a range of children from 3 to 8 years of age to allow more insight into the emergence of analogy over time. Second, previous work, including Carpentier et al., has primarily used MTS to train and test for derived relations. However, this procedure is often arduous and time-consuming, especially when young children are the participants; for example, Carpentier et al. (2003) reported that ‘subjects required 16 to 31 sessions of 16 to 22 min each over a period of 18 to 48 days’. In contrast, the present study employed an alternative format that allows much faster and more efficient training and testing of derived relations.

Previous RFT studies on assessing and training derived relations (i.e., AADRR) have successfully used the Relational Evaluation Procedure (REP) to do so. This methodology allows participants to report on or evaluate sets of arbitrarily applicable relations defined by various sets of contextual cues. In one of the most impressive examples of the applied educational utility of this technology, Cassidy, Roche, and Hayes (2011) and various follow-up studies (e.g., Hayes & Stewart, 2016; Cassidy et al., 2016) have used the REP to assess and train AADRR by presenting statements involving nonsense words juxtaposed with contextual cues (e.g., ‘CUG is the SAME as DAX’, ‘DAX is the SAME as YIM’) and asking participants (typically children from age 10 up to adolescence) questions that required them to derive relations based on those presented statements (e.g., ‘Is DAX the SAME as CUG?’, ‘Is CUG the same as YIM?’). Training with this methodology is extremely efficient as multiple exemplars of relational patterns can be easily generated and readily presented, and indeed training with this variation of the REP protocol has been shown to substantially boost children’s intellectual performance. Given that a key goal of the ARA protocol is the assessment of AADRR in terms of both basic relational framing itself as well as the relating of relations, the format employed drew on this type of REP format. At the same time, in accordance with the goals of the present study, the ARA protocol is required for the assessment of children much younger than Cassidy et al. (2011) or other studies that have adopted their methodology. Given the range of ages in the present study, it was highly probable that some participants, especially the younger ones, were not yet readers or had minimal reading skills. In order to allow for minimal reading skills and thus ensure that all participants could be assessed and trained equally effectively on AADRR, the stages in the ARA protocol focused on this repertoire utilised coloured circles, single letters, and audio options. It was envisaged that the use of this format would allow us to draw on the advantages of the REP (facilitating much more efficient and extensive testing of AADRR than matching to sample) with a much younger set of participants.

Third, Carpentier et al. focused on analogy and, more specifically, equivalence-equivalence responding alone. The training and testing format utilised in the current study was employed in the context of the multi-stage ARA protocol that allowed testing of a range of different types of relations (i.e., not just equivalence) and that did so at a number of different levels of complexity including (i) nonarbitrary relations, (ii) relating of nonarbitrary relations, (iii) arbitrary relations, and (iv) relating of arbitrary relations. This comprehensive protocol thus allowed for the assessment of analogical responding in the context of assessing participants' relational responding more broadly. In the present study, this protocol was used to assess participants' analogical relational performance and their relational performance more broadly and to correlate these with their intellectual performance as assessed on a standardised test of intellectual functioning, namely the Stanford-Binet Intelligence Scales (5^{th} Edition) for Early Childhood (SB5). This fed into a second broader aim of the present study, which was the assessment of the development of relational framing more generally in young children.

RFT has provided substantive evidence that AADRR underlies human language and cognition (Hayes et al., 2001), including work showing that training this repertoire can boost intellectual skill. However, despite the evidence that relational framing is a core skill underlying cognitive ability, research on the sequence of frame acquisition and the normative development of relational responding is as yet limited. Previous research on relational framing and age found that relational responding increased as a function of age in typically developing children. For example, McHugh et al. (2004) investigated the development of deictic frames in typically developing children and found that deictic responding improved with age. More recently, Mulhern et al. (2017) assessed containment and hierarchical relations in children and found a significant correlation between age and relational framing. In a series of studies, Dixon et al. (2014) and Dixon, Rowsey, et al. (2017) found strong correlations in both PEAK Phase 1 scores and age and PEAK Phase 2 scores and age. Hayes et al. (2001) proposed a hypothetical sequence for the development of increasingly complex AADRR but did not suggest a sequence of the emergence of specific frames.

In terms of the actual sequencing, there is limited data; the only other cross-sectional studies include the two previously mentioned studies by Mulhern et al. (2017) and McHugh et al. (2004). There is tentative evidence that suggests equivalence emerges at a young age (Lipkens et al.,1993), but on several other basic frames, there is very little research, and for distinction and temporality, for example, there is nothing. Barnes-Holmes, Barnes-Holmes, Smeets, et al. (2004) trained comparison in 4- to 6-year-old children, and Barnes-Holmes, Barnes-Holmes, and Smeets (2004) trained opposition in 4- to 6-year-old children, but no evidence was provided to suggest when these frames tend to emerge under normal conditions. The current cross-sectional study aimed to provide more comprehensive data on the emergence of basic patterns of framing across ages 3 to 7. Luciano et al. (2009) proposed a training sequence for early frames of coordination, opposition, distinction, comparison, and hierarchy and provided suggestions for intervention, including multiple exemplar training, bidirectional stimulus relations training, and systematically transitioning from training nonarbitrary stimulus relations to arbitrary relations. However, as yet, there is little or no empirical work on the normative acquisition of frames. This was the second aim of the current study in which we investigated the normative development of specific relational frames in young children and measured relational framing against standardised tests of cognitive abilities. This is one of the first attempts to provide a cross-sectional, comprehensive look at the acquisition of frames of coordination, comparison, opposition, temporality, and hierarchy in different age groups.

In the current study, we looked at relational framing across a number of different frames, including coordination, comparison, opposition, temporality, and hierarchy, across four different levels, including nonarbitrary relations, nonarbitrary analogy, arbitrary relations, and arbitrary analogy, and across a number of different age groups, specifically, three- to four-, four- to five-, five- to six-, and six- to seven-year-old children. One key purpose of the present study was to extend previous work on the development of relational framing broadly by examining the development of a range of different relational frames in young children. Previous RFT studies have used different methodologies to examine and train relational framing in young children. Arguably one of the most efficient methods is the relational evaluation procedure (REP). This methodology allows participants to report on or evaluate sets of arbitrarily applicable relations defined by various sets of contextual cues.

Given that a key goal of the present study was to assess relational framing across various levels, it was decided to employ a ‘relational statement’ format similar in some respects to that used by Cassidy et al. (2011). There was one critical difference, however. In Cassidy et al., participants were required to read sentences specifying and/or querying the relations between nonsense words (e.g., CUG is the same as DAX). Considering the range of ages in the present study (i.e., from 3 to 7 years), it seemed highly probable that some participants, especially the younger ones, were not yet readers or had minimal reading skills. Hence, to allow for minimal reading skills and thus ensure that all participants could be assessed and trained equally effectively on AADRR, the stages in the relational assessment focused on this repertoire utilised coloured shapes, single letters, and audio options. It was envisaged that using this format would allow us to draw on the advantages of the REP (facilitating much more efficient and extensive testing of AADRR than matching to sample) with a much younger set of participants.

The training and testing format utilised in the current study was employed in the context of the multi-stage protocol relational assessment that allowed testing of different types of relations. Also, in accordance with the more general aims of examining the development of relational framing more broadly, it did so at a number of different levels of complexity, including (i) nonarbitrary relations, (ii) relating of nonarbitrary relations, (iii) arbitrary relations, and (iv) relating of arbitrary relations. In addition, participants' relational performance across and within all levels and frames was correlated with their age and intellectual performance, as assessed on a standardised test of intellectual functioning, namely the Stanford-Binet Intelligence Scales (5th Edition) for Early Childhood (SB5).

Method

Participants and Settings

Participants included 24 students (14 females, 10 males) attending a private, non-secular American school in New Jersey. The demographic information for the participants is summarised in Table 3.1. Participant ages ranged from 36 to 84 months (*M* = 59.96, SD = 13.76), including five 3- to 4-year-old (*M* = 39.8 months, SD = 3.9), six 4- to 5-year-old (*M* = 53.33 months, SD 2.73), six 5- to 6-year-old (*M* = 65.67 months, SD 3.08), and seven 6- to 7-year-old participants (*M* = 75.14 months, SD 4.18). None of the participants had any known developmental or intellectual disabilities. Full-scale IQ scores, which were obtained for the purposes of this study, ranged from 94 to 132 (*M* = 113.25, SD = 10.59).

The assessment measures were administered by the researcher (the first author of the present study) at the participants' school in a separate, quiet classroom within the school building. When conducting the assessments, the researcher and the participant were seated at a child-sized school desk; the participant sat facing the desk, and the researcher sat at the side of the desk facing the participant and with a full view of the testing materials in front of the participant.

Table 3.1

Participant Demographic Information

		N	%	FSIQ	SD
Gender					
	Male	10	42		
	Female	14	58		
Age (months)					
	36-47	5	21	118.60	12.76
	48-59	6	25	110.83	9.00
	60-71	6	25	114.67	4.37
	72-84	7	29	110.29	14.01

Materials and Apparatus

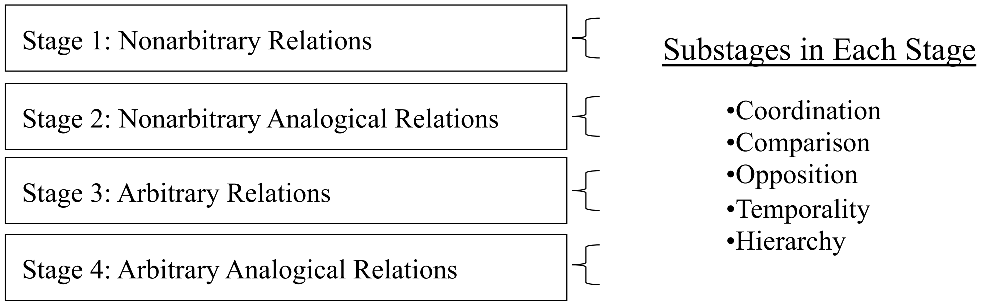
The materials used in this study included a standardised measure of intelligence, the Stanford-Binet Intelligence Scales – 5th Edition for Early Childhood (SB5) (Roid, 2003), and the relational assessment. The relational assessment was presented on a computer. Stages 1 and 2 of the assessment were presented on an Apple iPad Air 2 using Apple Keynote software, and Stages 3 and 4 were presented on an Acer Chromebook R 11 using Microsoft PowerPoint software.

The Stanford-Binet Intelligence Scales – 5th Edition for Early Childhood

The SB5 is an assessment of intelligence and cognitive abilities and is considered the standard measure of global intellectual ability for children and adults aged two to 85+ (Roid, 2003). In order to provide a reliable profile of differential abilities, the SB5 yields three composite scores, including full-scale IQ (FSIQ), Verbal IQ (VIQ), and Nonverbal IQ (NVIQ), and five-factor scores, including Fluid Reasoning (FR), Knowledge (KN), Quantitative Reasoning (QR), Visual-Spatial Processing (VS), and Working Memory (WM). Full-Scale IQ, VIQ, and NVIQ composite scores have high reliability coefficients ranging from .95 for VIQ to .98 for FSIQ (Roid, 2003). The reliability coefficients for substages range from .84 for Verbal-WM to .89 for Verbal-KN, and factor scores range from .90 for FR to .92 for KN, QR, and VS (Garred & Gilmore, 2009). The validity of the SB5 is supported based on its correlation with a number of alternative intelligence tests. For example, the SB5 has a significant correlation of .83 with the Wechsler Preschool and Primary Scale of Intelligence – Revised (WPPSI-R) and .78 with the Woodcock-Johnson III Tests of Cognitive Abilities – General Intellectual Ability – Standard (WJ III GIA). In the present study, to calculate correlations between the non-norm-referenced relational assessment scores and SB5 scores, raw scores were used for SB5 total, Verbal, and Nonverbal scores.

Figure 3.1

Schematic Representation of the Relational Assessment Stages and Substages



Note. This figure illustrates the elements and sequence of the relational assessment.

The Relational Assessment

The relational assessment examined different patterns of relational responding across four stages, including Stage 1: nonarbitrary (physical) relations; Stage 2: nonarbitrary analogical relations (relations between physical relations); Stage 3: arbitrary (abstract) relations (relational frames); and Stage 4: arbitrary analogical relations (relations between abstract relations). Within each stage were five substages focused on particular relations, including coordination, comparison, opposition, temporality, and hierarchy. Each substage included 10 trials (see Figure 3.1 for a schematic presentation of the relational assessment). In what follows, we present the general layout of the stimuli for each of the stages and substages in the assessment, and in the procedure section, we will describe the administration of the assessment. In Stages 1 and 2 of the relational assessment, the sample stimulus (in Stage 2, a compound stimulus) was presented at the top of the computer screen, and two comparison stimuli (in Stage 2, compound stimuli) were presented at the bottom left and right of the screen (for an illustrative example, see Appendices A1 and A2). In Stage 3, a relational network of stimuli was presented in the middle of the screen (see Appendix A3), while in Stage 4, a relational network of stimuli was presented at the top of the screen. A sample compound stimulus was in the middle of the screen, and the comparison compound stimuli were at the bottom left and right of the screen (see Appendix A4). For all trials, stimuli were delineated with a solid black line, including a horizontal black line underneath the sample and a vertical line between the comparison stimuli.

Stage 1: Nonarbitrary Relations. The relational stimuli included simple, monochromic pictures or shapes in the coordination, comparison, opposition, and temporal substages and boxes of varied sizes and colours in the hierarchy substage. The coordination substage included match-to-sample trials of a sample and three comparison stimuli. The comparison substage included trials in which the participant was shown a sample and three comparisons; the comparisons were identical to the sample except for size and included one bigger, one smaller, and one identical comparison. The opposition substage included trials in which the participant was shown a sample and three comparisons identical to the sample except for variance on one dimension on a gradient scale. For example, if the sample was a black square, the comparisons included a white, grey, and black square. The temporality substage included trials in which the participant was instructed to watch the iPad screen for sequentially appearing shapes. The first stimulus appeared 0.5 s after the onset of a new trial, and the second stimulus appeared 1.0 s after the first stimulus. The hierarchy substage included trials in which the participant was shown two or three different sized and different coloured boxes.

Stage 2: Nonarbitrary Analogical Relations. In Stage 2, all stimuli presented were compound stimuli. A sample compound stimulus was presented at the top of the screen in all substages, and two comparison compound stimuli were presented below the sample. The stimuli used in the sample compound were always different from the stimuli used in the comparison compounds. In the coordination substage, the sample was composed of either two identical or two nonidentical pictures, and the comparisons included compound stimuli composed of identical stimuli and nonidentical stimuli. In the comparison substage, each of the sample and comparison compounds included three stimuli that were similar except for size. Two of the three stimuli in each compound were outlined in red, and an arrow indicated which stimulus to compare with another stimulus in the compound (i.e., from smaller to bigger or bigger to smaller). In the opposition substage, each sample and comparison compound included three stimuli that varied on a gradient scale along one dimension (e.g., size, colour, quantity). In each compound, two of the stimuli had a solid red line underneath and were either nonarbitrarily opposite or not opposite. In the temporality substage, both the sample and comparison compounds included two stimuli that appeared onscreen either simultaneously or 0.5 seconds apart. In the hierarchy substage, each trial included sample and comparison stimuli consisting of a square inside another square and small, blue dots located in either the innermost or outermost square or outside the squares.

Stage 3: Arbitrary Relations. An adaptation of the REP format was employed in all substages. The relational stimuli were simple, monochromic shapes (circles, triangles, squares) separated by a single Latin letter indicating the contextual cue (S for Same, D for Different, M for More, L for Less, O for Opposite, B for Before, A for After, C for Contains, and I for Inside), plus corresponding audio icons (which, when touched, produced an audio recording of the contextual cue (e.g., 'same') through the laptop or computer speaker).

Stage 4: Arbitrary Analogical Relations. The REP format was also used in Stage 4, including the use of monochromic circles for the relational stimuli plus visual and audio signals representing the contextual cues. The sample and comparison compounds were combinations of the relata in the relational network.

Procedure

At the start of the session, the researcher placed the iPad or computer in front of the participant. Participants were provided with detailed instructions before they commenced with each substage in every stage. Responses were scored as either correct or incorrect. A correct response was defined as touching or pointing to the correct comparison or providing a correct vocal response. An incorrect response was defined as touching or pointing to the incorrect comparison, touching both correct and incorrect comparisons, providing an incorrect vocal response, not responding, or engaging in other behaviour that could not be categorised as correct. Participants did not receive feedback for any responses during the relational assessment or SB5 testing. Generalised reinforcement was provided for compliance and participation throughout the assessments.

The sequence in which the two assessments were administered was randomised so that participants completed either the relational assessment or the SB5 first. The SB5 was completed in one session, and the relational assessment was completed in one to two sessions. For the relational assessment, each session lasted between 20 and 60 min, including scheduled breaks every 10-15 min, or if the participant requested a break. For all participants, both assessments were conducted within a maximum of four weeks of each other.

Prior to conducting the study, ethical approval for the recruitment of participants was obtained from the research ethics committee of the host (research) institution. Consent for conducting the study was obtained from the principal of the New Jersey school. Caregiver consent was obtained for each child who participated, and verbal consent was also obtained from each participant.

Stage 1: Nonarbitrary Relations

In the coordination substage, participants were asked to match the sample to a comparison upon hearing the instruction, ‘Which one of these at the bottom is like this one at the top?’ In the comparison substage, participants were asked to identify the correct comparison in relation to the sample upon hearing the instruction, ‘Which one of these is *bigger* or *smaller* than the one at the top?’ In the opposition substage, participants were asked to identify the correct comparison in relation to the sample upon hearing the instruction, ‘Which one of these is *opposite* to the one at the top?’ In the temporality substage, participants were to identify the order in which the stimuli appeared upon hearing instructions such as, ‘Which one was *before* or *after* the other one/was stimulus 1 before or after stimulus 2?’ In the hierarchy substage, participants were to identify where a box was in relation to another box upon hearing instructions such as, ‘Which box is inside/which box *contains* the other box?’

Stage 2: Nonarbitrary Analogical Relations

In Stage 2, all stimuli presented were compound stimuli; in all substages, the participants’ task was to match the sample (compound stimulus) relation to the correct comparison (compound stimulus) relation. For example, in the coordination substage, the sample was composed of either two identical or two nonidentical pictures, and the comparisons included a compound stimulus composed of two identical stimuli and a compound composed of two nonidentical stimuli. Participants were required to match the sample to one of the comparisons upon hearing the instruction, ‘Look at these at the top. Which one of these (point to the comparison stimuli) is like the ones at the top?’ A pre-trial sample was presented before administering each substage in order to familiarise the participant with the testing format (for an illustrative example, see Appendix A2).

Stage 3: Arbitrary Relations

Prior to starting Stage 3, a pre-test was administered to familiarise participants with the test format. The participant was shown a computer screen with a relational network, for example [Red Circle] [S] [Blue Circle]. The assessor instructed the participant to look at the screen and said, ‘The S means same. If you can’t remember what the S means, you can tap on the S, and the computer will tell you.' Once the participant was comfortable with the visual and audio stimuli, the assessor provided the instruction, 'Let's start. We are going to pretend these shapes like food, and we are going to talk about whether they like or do not like the same food. Look at the screen in front of you.' A relational network was presented on the screen, for example [Red Circle] [S] [Blue Circle]; [Blue Circle] [S] [Yellow Circle], and the assessor read, 'Red likes the same food as Blue, and Blue likes the same food as Yellow.' The assessor asked questions about the relational network, including questions about directly given relations (e.g., 'Does Red like the same food as Blue?'), questions requiring mutual entailment (e.g., 'Does Red like different food to Blue?'), and questions requiring combinatorial entailment (e.g., 'Does Red like the same food as Yellow?'). The pre-test included ten yes/no questions.

For all substages in Stage 3, the assessor first read the relational network to the participant and then asked the trial questions. Questions became increasingly difficult and required responses, including directly trained, mutually entailed, and combinatorially entailed relations. The coordination substage was like the pre-test but introduced novel stimuli.
In the comparison substage, the relational network included coloured circles and the contextual cues more (M) and less (L). Participants were told that they were going to pretend-buy their favourite food or candy, 'These circles are like coins, and we are going to pretend that you can use them to buy your favourite food/candy.'[1] For example, the participant saw: [Blue Circle] [M] [Red Circle], and the assessor said: 'Blue buys more than Red, so which coin should you take to the store to buy [insert participant's favourite food]?' The correct choice, in this case, would have been to select the blue coin because blue buys more.

In the opposition substage, the relational network included the contextual cue (O) for opposite and coloured circles representing coins. Participants were told that they were going to pretend-buy their favourite food or candy: 'These circles are like coins, and we are going to pretend that you can use them to buy your favourite food/candy.' For example, the participant saw: [Red Circle] [O] [Blue Circle], and the assessor said, 'Red buys many/a lot/a few/a little. Red is opposite to Blue, which coin should you take to the store to buy [insert participant's favourite food]?' The correct choice on every trial was the selection of the colour worth a lot/many.

In the temporal substage, the relational network included coloured squares and the contextual cues *before* (B) and *after* (A). The assessor told the participant that the coloured

---

[1] Adaptation of the procedure used in Barnes-Holmes, Barnes-Holmes, Smeets, Strand, and Friman (2004).

squares were in a race: ‘These silly little squares are racing. Which one reaches the finish line before/after the other one?’ For example, the participant saw: [Red Square] [B] [Blue Square], and the assessor said, ‘Red was before Blue. Was Blue before/after Red?’

In the hierarchy substage, the relational network included coloured circles and the contextual cues *inside* (I) and *contains* (C). The assessor told the participant that the coloured circles either contained each other or were inside each other. For example, the participant saw: [Red Circle] [I] [Blue Circle], and the assessor said, ‘Red is inside Blue.’ The assessor would then present one of the following questions depending on the particular trial type: ‘Is Blue inside Red? Does Blue contain Red? Does Red contain Blue? Which one is inside? Which one contains the other one?’

Stage 4: Arbitrary Analogical Relations

The same relational network composed of circles and single letters was used for all trials in each substage. The sample and comparison compounds were combinations of the relata in the relational network. For example, in the coordination substage, the relational network would be read as, ‘Blue is the same as yellow, yellow is the same as red, and red is different to green.’ The sample compound may be [Blue : Red], and the comparison compounds may be [Yellow : Green] and [Red : Yellow] (see Appendix A4 for an illustrative example of each substage). On each trial, the researcher read the relational network to the participant and then 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?’ For example, given the relational network mentioned above, [Blue S Yellow], [Yellow S Red], and [Red D Green], participants could derive the relation of sameness for a sample stimulus [Blue : Red], a relation of sameness for a comparison stimulus [Red : Yellow], and a relation of difference for a comparison stimulus [Yellow : Green]. This example was for the coordination substage; for each of the other substages, the participants had to match the correct comparison relation to the appropriate sample relation, but the nature of the relations that had to be matched depended on which substage it was. There were ten trials in each substage, including trials for directly trained, mutually entailed, and combinatorially entailed relations in that order (i.e., increasing difficulty).

Interobserver Agreement

Interobserver agreement was calculated for 20% of all completed relational assessment substages by a trained research assistant. The research assistant was trained in data collection until they reached 100% accuracy prior to collecting IOA data. Trial-by-trial IOA was calculated and ranged from 97% to 100% (M = 98.66%). Agreement across all measures was counted if both observers provided the same score for each item assessed, and disagreement was counted if one observer provided a different score compared to the other observer. Interobserver agreement was calculated for 24% of the verbal (Vocabulary) and nonverbal (Object Series/Matrices) routing substages of the SB5. Trial-by-trial IOA was calculated and ranged from 94% to 100% (M = 97.83%). Treatment integrity was evaluated on 20% of all completed relational assessment substages and 23.5% of all SB5 routing substages by a trained research assistant. The observer scored a + on each trial that the trainer was 1) observed to gain attention prior to the trial, 2) accurately presented the question, and 3) consequated appropriately. Mean procedural integrity ranged from 97% to 100% (M = 99%) for the relational assessment substages, and from 88% to 100% (M = 98.1%) for the SB5 routing substages.

Results

In the results section, we will examine the relational assessment protocol scores, correlations between relational assessment scores and age, correlations between relational assessment scores and raw IQ, correlations within the protocol, and finally, the acquisition of relations across age groups. Results for Stages 2 (nonarbitrary analogy) and 4 (arbitrary analogy) will receive particular attention in each subsection. In addition, because the percentage of correct responding is provided as a performance outcome across different trial types, in Appendix B, we provide a table showing the chance level of responding (based on a possible number of outcomes) for different trial types. In the section below, in which we examine performance across relational frames, we compare the performance of the different cohorts on different stages and for different relations with chance-level performance.

Relational Assessment Protocol Scores

Table 3.2 shows scores on both the relational assessment (overall and by stage) and the SB5 (full scale and subscale) both as a function of the entire cohort as well as by age group. Total relational assessment scores across all participants ranged from 26.5% to 78% correct (*M* = 56%, SD 28.58). Stage 1 (nonarbitrary relations) scores across all substages and all participants ranged from 58% to 94% (*M* = 80%, SD 5.19); Stage 2 (nonarbitrary analogy) scores ranged from 40% to 84% (*M* = 61%, SD 5.46); Stage 3 (arbitrary relations) scores ranged from 0% to 88% (*M* = 54%, SD 10.70); and Stage 4 (arbitrary analogy) scores ranged from 0% to 66% (*M* = 29%, SD 12.53).

Table 3.2

Participant Scores on the Relational Assessment and SB5

ID	M/F	Age (m)	SB5- FSIQ	SB5- NV	SB5- V	SB5- T	RA- T	RA1 NAR	RA2 NAA	RA3 AR	RA4 AA
3- to 4-year-old cohort											
S4	F	36	127	46	57	103	62	29	24	9	0
S26	F	37	132	54	59	113	104	35	26	17	26
S3	F	40	100	37	40	77	53	33	20	0	0
S5	F	40	122	55	48	103	77	39	28	10	0
S17	F	46	112	55	50	105	84	43	27	14	0
Mean		39.80	118.60	49.40	50.80	100.20	76.00	35.80	25.00	10.00	5.20
SD		3.90	12.76	7.89	7.60	13.61	19.84	5.40	3.16	6.44	11.63
4- to 5-year-old cohort											
S11	M	52	110	59	60	119	80	32	26	22	0
S22	M	55	102	50	62	112	102	43	26	29	4
S23	M	49	126	62	68	130	116	42	34	31	9
S27	F	53	103	51	57	108	93	37	24	28	4
S34	F	54	108	54	66	120	95	38	27	30	0
S60	F	57	116	74	64	138	90	31	24	28	7
Mean		53.33	110.83	58.33	62.83	121.17	96.00	37.17	26.83	28.00	4.00
SD		2.73	9.00	8.96	4.02	111.18	12.15	4.96	3.71	3.16	3.63
5- to 6-year-old cohort											
S39	M	68	118	82	80	162	134	41	42	34	17
S40	F	61	110	64	76	140	105	38	32	31	4
S46	F	63	121	78	76	154	143	46	37	27	33
S47	M	67	115	79	69	148	101	41	31	25	4
S49	M	66	110	73	67	140	115	43	29	22	21
S59	M	69	114	81	77	158	127	46	32	26	23
Mean		65.67	114.67	76.17	74.17	150.33	120.83	42.50	33.83	27.50	17.00
SD		3.08	4.37	6.74	5.04	9.24	16.62	3.15	4.79	4.32	11.37
6- to 7-year-old cohort											
S52	F	75	94	76	69	145	136	40	32	27	27
S53	M	76	123	87	101	188	156	47	40	41	28
S54	M	72	128	86	99	185	138	37	33	35	33
S55	F	74	100	74	70	144	138	42	32	38	26
S58	F	72	95	74	63	137	136	47	35	35	19
S56	M	73	120	80	92	172	144	43	30	44	27
S61	F	84	112	93	92	185	146	46	37	33	30
Mean		75.14	110.29	81.43	83.71	165.14	142.00	43.14	34.14	37.57	27.14
SD		4.18	14.01	7.39	15.83	22.37	7.30	3.80	3.44	3.82	4.30
All participants											
Mean		59.96	113.25	67.67	69.25	136.92	111.46	39.96	30.33	26.92	14.25
SD		13.76	10.59	14.97	15.46	29.25	28.58	5.19	5.46	10.70	12.53

Note. SB5-FSIQ = Stanford-Binet Intelligence Scales – 5th Edition for Early Childhood Full-Scale IQ; NV = nonverbal; V = verbal; T = total; RA = relational assessment; NAR = nonarbitrary relations; NAA = nonarbitrary analogical relations; AR = arbitrary relations; AA = arbitrary analogical relations.

Relational Assessment Protocol Scores and Age

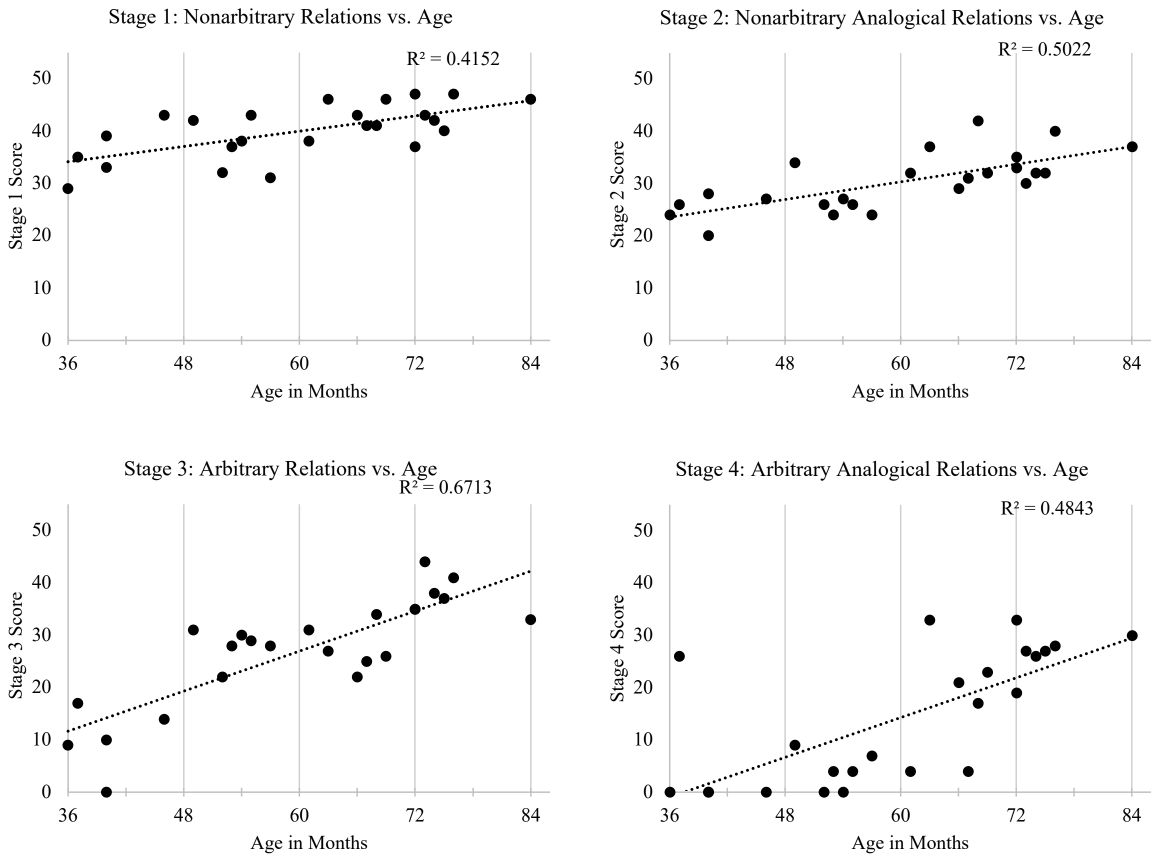
Concerning breakdown by age group, we can see relatively predictable patterns wherein relational assessment scores increased as a function of age and mean relational assessment scores increased by approximately 10% for each additional year in age. Specifically, total relational assessment scores ranged from 26.5% to 52% correct (M = 38%, SD 19.84) for the 3- to 4-year-old cohort, from 40% to 58% correct (M = 48%, SD 12.15) for the 4- to 5-year-old cohort, from 50.5% to 71.5% correct (M = 60.4%, SD 16.62) for the 5- to 6-year-old cohort, and from 68% to 78% correct (M = 71%, SD 7.3) for the 6- to 7-year-old cohort (refer to Appendix B for details on chance-level responding).

Mean scores within each stage showed improvements across age cohorts. The biggest improvement in relational assessment scores based on age occurred in the later stages, Stages 3 (arbitrary relations) and 4 (arbitrary analogy). For example, comparing the youngest with the oldest cohorts, we see that in Stage 3, the 3- to 4-year-old cohort scored on average 20% correct across all substages while the 6- to 7-year-old cohort scored 75% correct, a difference of 55%. In Stage 4, the 3- to 4-year-old cohort scored on average 10% correct while the 6- to 7-year-old cohort scored 54% correct, a difference of 44%. In contrast, in Stages 1 and 2, there was less of an age-based gap in performance. In Stage 1, the 3- to 4-year-old cohort scored an average of 72% correct while the 6- to 7-year-old cohort scored 86% correct, a difference of only 14%, while in Stage 2, the 3- to 4-year-old cohort scored an average of 50% correct while the 6- to 7-year-old cohort scored 68% correct, a difference of 18%.

Apart from considering the general improvement based on age, we can also examine each stage to check for particular discontinuities between specific age cohorts. For example, in Stage 2 (nonarbitrary analogy), there is a large gap in performance between the 4- to 5-year-old cohort and the 5- to 6-year-old cohorts (*M* = 54% to 68%, respectively). These data show that the transition from the 4- to 5-year-old to the 5- to 6-year-old age range results in a particular improvement in nonarbitrary analogy. In Stage 3 (arbitrary relations), the biggest performance improvement occurred between the 3- to 4-year-old and 4- to 5-year-old cohorts (*M* = 20% to 56%, respectively). These data thus suggest substantial development in arbitrary relational framing between these ages. In Stage 4 (arbitrary analogy), the biggest improvement in performance occurred between the 4- to 5-year-old and 5- to 6-year-old cohorts (*M* = 8% and 34%, respectively). This finding for arbitrary analogy is similar to that for nonarbitrary analogy in Stage 2. The fact that there appears to be a substantive improvement in analogical ability between the 4- to 5-year-old and 5- to 6-year-old participants at both the nonarbitrary and arbitrary levels supports earlier research findings (e.g., Carpentier et al., 2002) that analogical ability begins to emerge around five years of age. One final point concerns the difference in variability in performance in the arbitrary analogy between different cohorts. In particular, the standard deviation decreased from 11.37 for the participants aged 5 to 6 years old to 4.3 for participants aged 6 to 7 years old. This suggests less variable, more stable responding in the older cohort following the initial emergence of analogy.

Figure 3.2

Relational Assessment vs. Age Across Stages



Note. Regression slopes are shown for participant scores at each stage and age.

Table 3.3

Matrix of Spearman's Rho Correlations for All Measures Administered

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1. Age	-																												
2. SB5-T	.879**	-																											
3. SB5-NV	.856**	.958**	-																										
4. SB5-V	.832**	.973**	.894**	-																									
5. RA-T	.859**	.874**	.810**	.871**	-																								
6. S1-T	.604**	.512*	.534**	.484*	.688**	-																							
7. S2-T	.709**	.782**	.806**	.788**	.831**	.684**	-																						
8. S3-T	.818**	.717**	.637**	.753**	.817**	.430*	.649**	-																					
9. S4-T	.746**	.815**	.760**	.774**	.915**	.504*	.672**	.639**	-																				
Stage 1 Substages: 1 = Coordination; 2 = Comparison, 3 = Opposition; 4 = Temporality; 5 = Hierarchy																													
10. SS-1	0.316	0.349	0.327	0.371	.436*	0.394	0.394	0.381	0.396	-																			
11. SS-2	.548**	.595**	.578**	.534**	.625**	.516**	.477*	.439*	.631**	.416*	-																		
12. SS-3	0.332	0.225	0.242	0.206	.429*	.821**	.459*	0.28	0.224	0.265	.474*	-																	
13. SS-4	0.393	0.313	0.355	0.277	0.357	.528**	0.252	0.122	0.321	0.121	-0.084	0.178	-																
14. SS-5	0.384	.431*	.430*	.487*	.577**	.564**	.767**	.457*	.470*	.483*	0.165	0.294	0.211	-															
Stage 2 Substages: 1 = Coordination; 2 = Comparison, 3 = Opposition; 4 = Temporality; 5 = Hierarchy																													
15. SS-1	.605**	.702**	.693**	.680**	.648**	.422*	.714**	.429*	.657**	0.333	.590**	0.151	0.212	.475*	-														
16. SS-2	0.326	.499*	.469*	.541**	.477*	.558**	.592**	0.318	0.31	0.269	0.299	.434*	0.156	.527**	0.235	-													
17. SS-3	0.211	0.289	0.295	0.318	0.288	.446*	.556**	0.154	0.085	0.238	-0.125	0.322	0.259	.563**	0.17	0.343	-												
18. SS-4	0.329	0.154	0.171	0.134	0.262	0.389	0.355	0.226	0.114	0.123	0.199	.418*	0.069	0.262	0.089	-0.041	0.335	-											
19. SS-5	.568**	.582**	.637**	.557**	.683**	0.374	.690**	.641**	.657**	0.154	0.376	0.173	0.185	.456*	.612**	0.221	0.049	-0.047	-										
Stage 3 Substages: 1 = Coordination; 2 = Comparison, 3 = Opposition; 4 = Temporality; 5 = Hierarchy																													
20. SS-1	.658**	.605**	.555**	.590**	.682**	.643**	.621**	.687**	.521**	.419*	.526**	.625**	0.121	.432*	.427*	0.346	0.359	0.237	.551**	-									
21. SS-2	.730**	.715**	.631**	.693**	.837**	.421*	.525**	.754**	.826**	.419*	.687**	0.175	0.135	0.286	.474*	0.217	0.008	0.196	.513*	.455*	-								
22. SS-3	.568**	.455*	.458*	.470*	.554**	0.298	.579**	.718**	.428*	0.274	0.17	0.116	-0.004	.571**	0.312	0.242	0.338	0.246	.493*	0.362	.533**	-							
23. SS-4	.676**	.684**	.606**	.704**	.632**	0.185	.546**	.781**	.480*	0.033	0.209	0.031	0.065	0.28	0.356	0.392	0.144	0.101	.530**	.416*	.506*	.569**	-						
24. SS-5	.640**	.560**	.436*	.620**	.567**	0.229	0.31	.812**	.442*	0.378	0.345	0.194	0.147	0.19	0.318	0.098	-0.083	0.122	0.353	.586**	.512*	0.336	584**	-					
Stage 4 Substages: 1 = Coordination; 2 = Comparison, 3 = Opposition; 4 = Temporality; 5 = Hierarchy																													
25. SS-1	.703**	.749**	.667**	.729**	.829**	.551**	.548**	.557**	.865**	0.399	.496*	0.262	.453*	0.347	.576**	0.341	0.114	-0.01	.530**	.444*	.741**	0.375	.450*	.432*	-				
26. SS-2	.702**	.685**	.678**	.618**	.850**	.579**	.661*	.595**	.885**	0.339	.703**	.409*	0.23	0.348	.612**	0.273	0.055	0.124	.731**	.629**	.734**	.405*	.428*	0.315	.751**	-			
27. SS-3	.657**	.776**	.759**	.737**	.864**	.445*	.737**	.613**	.911**	0.317	.546**	0.221	0.222	.458*	.610**	0.403	0.146	0.025	.759**	.489*	.774**	.468*	.480*	0.305	.782**	.837**	-		
28. SS-4	.648**	.658**	.616**	.598**	.769**	0.403	.491*	.476*	.893**	0.262	.599**	0.158	0.296	0.289	.501*	0.127	-0.073	0.22	.505*	0.392	.738**	0.184	0.306	0.356	.708**	.792**	.768**	-	
29. SS-5	.583**	.620**	.573**	.560**	.747**	0.389	.486**	0.403	.894**	0.262	.599**	0.14	0.278	0.308	.529**	0.132	-0.059	0.225	.479*	0.322	.733**	0.172	0.258	0.263	.726**	.779**	.782**	982**	-

Note. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed); SB5 = Stanford-Binet Intelligence Scales – 5th Edition for Early Childhood Full-Scale IQ; NV = nonverbal; V = verbal; T = total; RA = relational assessment; S- = stage; SS- = substages.

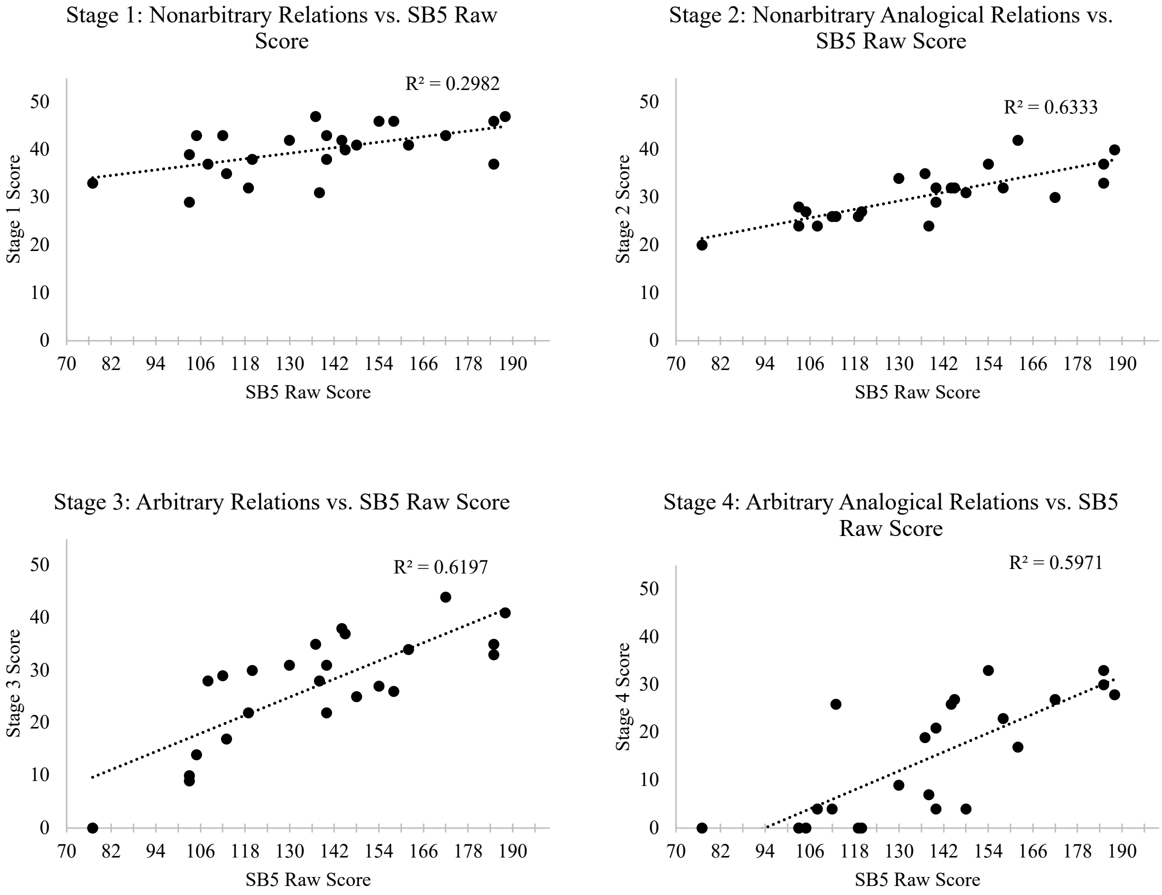
Correlating Assessment Protocol Performance and Age

Figure 3.2 shows graphs of the relationships between scores on each of the four relational assessment stages and age. These data suggest strong correlations between each of the stages and age. Table 3.3 shows a matrix of Spearman's rank correlations between variables, including age, SB5 total, and substage raw scores, and relational assessment total and substage scores. There is a strong correlation between relational assessment total score and age ($r = .859$, $p < 0.01$). Further analyses also show strong, significant correlations between total scores for each stage and age (Stage 1: $r = .604$, $p < 0.01$; Stage 2: $r = .709$, $p < 0.01$; Stage 3: $r = .818$, $p < 0.01$; and Stage 4: $r = .746$, $p < 0.01$). All correlations across stages and age were significant at the .01 level. These correlations suggest, consistent with RFT, that the capacities to engage in relational responding and analogical responding (both overall as well as across different frames) are established and strengthened via ongoing exposure to the typical socio-verbal environment.

Within-stage analyses show a significant correlation between age and comparison relations in Stage 1, and age and coordination, and hierarchy in Stage 2. These analyses show significant correlations between age and all substages in Stages 3 and 4.

Figure 3.3

Relational Assessment Scores vs. Raw IQ Score Across Stages



Note. Regression slopes are shown for participant scores at each stage and SB5 raw scores.

Relational Assessment Protocol Performance and Raw IQ

Figure 3.3 shows graphs of the relationships between scores on each of the four relational assessment substages and raw IQ scores. Spearman's rank analysis reveals a strong correlation between relational assessment total scores and raw IQ scores ($r = .874$, $p < 0.01$). Further analyses show strong correlations between each stage of the relational assessment scores and raw IQ scores (Stage 1: $r = .512$, $p < 0.05$; Stage 2: $r = .782$, $p < 0.01$; Stage 3: $r = .717$, $p < 0.01$; and Stage 4: $r = .815$, $p < 0.01$). Furthermore, each stage of the relational assessment also correlated with both SB5-NV and SB5-V subscale raw scores (SB5-NV and Stage 1: $r = .534$, $p < 0.01$; Stage 2: $r = .806$, $p < 0.01$; Stage 3: $r = .637$, $p < 0.01$; and Stage 4: $r = .760$, $p < 0.01$; and SB5-V and Stage 1: $r = .484$, $p < 0.05$; Stage 2: $r = .788$, $p < 0.01$; Stage 3: $r = .753$, $p < 0.01$; and Stage 4: $r = .774$, $p < 0.01$). Altogether, these correlations suggest, consistent with RFT, a strong link between relational responding (both overall as well as across different frames) and intellectual potential, as reflected in the raw IQ score.

Within-stage analyses show a significant correlation between Stage 1 comparison and hierarchy scores and SB5 raw score. These findings are consistent with previous research showing the importance of comparative relational framing for intellectual potential (Cassidy et al., 2011, 2016). There are also significant correlations between Stage 2 coordination, comparison, and hierarchy scores, and SB5 raw score. Finally, in Stages 3 and 4, there are significant correlations between all substage scores and SB5 raw scores. This pattern is consistent with the idea that arbitrary relational framing is particularly important for intellectual potential. Upon further examination across all four stages, comparison and hierarchy are the only substages that show a consistent correlation with raw IQ as revealed by both SB5 total score as well as SB5 subscale scores. This pattern suggests that comparison and hierarchy may be particularly important foundational intellectual skills.

Correlating Intra-Protocol Relations

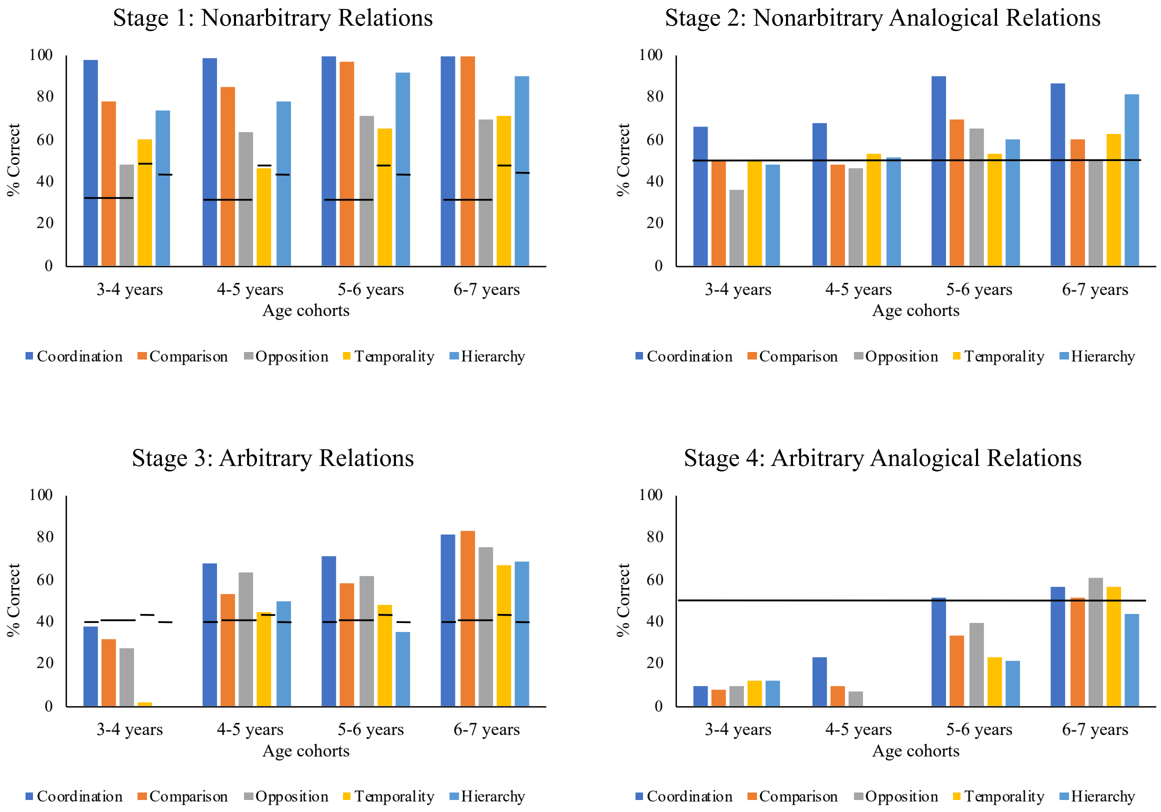
Spearman's rank analysis in Table 3.3 shows strong, significant relationships within and across many of the stages and substages. Stage 1 comparison and hierarchy scores show significant correlations across substage total scores for all stages; these data may suggest that nonarbitrary comparison and hierarchy are important prerequisite relations for more complex language. Stage 2 coordination scores show significant correlations across substage total scores for all stages; these data indicate that performance on nonarbitrary coordinate analogy may be a good predictor of relational responding more generally. Stage 3 coordination and comparison scores show significant correlations across substage total scores for all stages, while Stage 4 coordination, comparison, and opposition scores show significant correlations across substage total scores for all stages. Again, this seems to indicate that arbitrary relational framing may be particularly important for intellectual potential in general.

Only the comparison substage in Stage 1 shows significant correlations with age, IQ, and total relational assessment score. The coordination and hierarchy substages in Stage 2 show significant correlations with age, IQ, and total relational assessment score. All Stage 3 and Stage 4 substages show strong, significant correlations with age, raw IQ score, and total relational assessment score. These findings further support the key importance of arbitrary relational responding (i.e., relational framing) as a core intellectual ability acquired in childhood.

In both Stages 1 and 3, the comparison substage is strongly correlated with all of the substages in Stage 4. These data suggest nonarbitrary and arbitrary comparative relational repertoires are important intellectual skills and are strong predictors of analogical ability. Stage 1 comparison scores also show a significant correlation with Stage 2 coordination scores, which in turn, as previously noted, shows significant correlations with age, IQ, and total relational assessment score. Thus, it may seem that having a strong nonarbitrary comparative repertoire facilitates nonarbitrary coordinate analogy, which supports nonarbitrary and arbitrary relational responding.

Figure 3.4

Mean Normative Acquisition Across Stages and Relations



Stage 1: Nonarbitrary Relations

Stage 2: Nonarbitrary Analogical Relations

Stage 3: Arbitrary Relations

Stage 4: Arbitrary Analogical Relations

Note. Mean relational assessment scores across substages and age cohorts for each stage in the relational assessment. The black bars represent chance-level responding for each frame across each level.

Examining Performance Across Relational Frames

One final analysis looked at the acquisition of relational framing in each stage. Figure 4 shows the mean percent correct for each relational frame across stages and for each age cohort.

In general, in Stage 1 (nonarbitrary relations), there is an improvement with age across cohorts. Performance across cohorts shows very little variation in coordination (mean scores range from 98% to100% correct). Comparative and hierarchical relational performance start to emerge and show steady improvement as age increases, especially between the 4- to 5-year-old and 5- to 6-year-old cohorts. Participants across all cohorts perform lower in temporality and opposition, and performance in both these frames gradually improves across cohorts. Comparing performance across Stage 1 relations with chance-level responding (see Figure 4), we can see that the two older cohorts respond above chance levels on all relations while the two younger cohorts perform above chance levels in all relations except temporality.

In Stages 2 (nonarbitrary analogy) and 3 (arbitrary relations), performance across all relational frames generally improves with age. The most obvious improvement across all frames with age occurs in Stage 3. Specifically, there is a very clear increase in performance between the 3- to 4-year-old and 4- to 5-year-old cohorts. Comparing performance across Stage 2 relations with chance-level responding, we can see that the only relation on which all cohorts are responding above chance is sameness. In the case of other relations, the two younger cohorts are at or below chance in the case of all relations[2], while the two older cohorts both show above-chance-level responding for all other relations except temporality (5- to 6-year-olds) and opposition (6- to 7-year-olds). By showing how, in a slightly unusual context (i.e., nonarbitrary analogy), even older children can still show weak performance on these relational repertoires, the latter pattern perhaps suggests their potential difficulty. Comparing performance across Stage 3 relations with chance-level responding, we note first that the youngest group fails to meet a criterion of chance level of responding on any relations while the oldest group does so for all relations. The two other groups are somewhere in between, with both showing strong performance relative to chance for sameness,

---

[2] Regarding participant behaviour in low scoring trials: All cohorts continued to respond to questions, stay on task, and provide answers throughout assessment. In cases where the data show that a cohort was responding below chance levels, participants were generally responding with “I don’t know” or incorrect responses, especially as the questions became more difficult.

comparison, and opposition while showing near chance-level responding for both temporality and hierarchy.

As regards analogy in Stage 4 (arbitrary analogy), the 3- to 4-year-old and 4- to 5-year-old cohort scores are quite low and well below chance-level performance. Performance shows substantial improvement in the 5- to 6-year-old cohort, and this improvement continues in the 6- to 7-year-old cohort. Furthermore, in the latter group, analogical performance is improved not just in coordination but also in various other frames. Specifically, the 5- to 6-year-old cohort shows obvious improvement in performance in coordinate analogy. However, participants are still not doing that well (even the coordination performance is only just above chance level), indicating that analogy emerges at this age, but it is still a fragile repertoire. The data show gradual improvement by the 6- to 7-year-old cohort, and frames beyond coordination are emerging and improving. It might be noted from the graph that even the oldest cohort is still not responding much above the chance level for any of the relations. However, a more detailed analysis of our data in this respect showed that children in the group were not responding randomly on trials as they tended to answer simpler trials correctly while failing more difficult ones. This pattern was also observed more generally across all stages and all relations, suggesting that the comparison with chance-level responding should be seen as a guide rather than indicative of whether children were actually responding at random or not.

Summary

The normative data suggest, consistent with RFT, that the capacity to engage in relational responding and analogical responding (both overall as well as across different frames) are established and strengthened via ongoing exposure to the typical socio-verbal environment. Relational performance is correlated with measured IQ, which supports the RFT concept that relational framing is critical to language and cognition. Furthermore, the arbitrary stages are more highly correlated with IQ than the nonarbitrary stages. Comparison framing seems to be a particularly important relation as it shows significant correlations with age, IQ, and total relational assessment score. Considering the five relational frames tested in the assessment, several patterns are evident. Nonarbitrary and arbitrary coordination emerges first, and temporality emerges last. Also, there is a difference between nonarbitrary and arbitrary relational responding for frames of opposition and hierarchy; opposition scores are lower in the nonarbitrary stages, and higher in the arbitrary stages, and the hierarchy scores are higher in the nonarbitrary stages but lowest in the arbitrary stages.

Discussion

This study is one of the first attempts to provide a cross-sectional assessment of the acquisition of analogical responding, a particularly important form of relational framing with respect to intellectual potential (Sternberg, 1977; Sternberg et al., 2001). Analogical responses across frames of coordination, comparison, opposition, temporality, and hierarchy were assessed in 3- to 7-year-old children at both non-arbitrary and arbitrary levels. Performance was correlated with age and intellectual skill as assessed on a standardised IQ test. In addition, a second broader aim of this study was to measure the acquisition in young children of relational responding more generally. Accordingly this study also measured young children's relational responding more broadly across the various frames already listed and once again at both non-arbitrary and arbitrary levels.

Analogical Responding

The present study assessed relating relations (i.e., analogy) in young children at both nonarbitrary and arbitrary relational levels. It is commonly accepted in mainstream cognitive research (e.g., Bod, 2009; Gentner, 1983) that analogy is a particularly important pattern of relational framing with regard to intellectual potential. Moreover, analogical reasoning is frequently applied as a metric of intelligence (Sternberg, 1977) and as a measure to predict academic success; for example, the Law School Admissions Test (LSAT) (Lapiana, 2004). The theoretical importance of analogy has also been recognised explicitly within the RFT literature. For example, the relating of relations is included as one of the five levels (including mutual entailing, relational framing, relational networking, relating relations, and relating relational networks) in Barnes-Holmes et al.'s (2017, 2020) recently proposed hyper-level multi-dimensional (HDML) framework for examining and discussing relational framing. However, despite the ubiquity and utility of analogy, functional assessments targeting relations among relations as analogy have not yet been examined.

The present data extend previous work on the assessment of analogy by Carpentier et al. (2002, 2003, 2004). However, the present study arguably advances the Carpentier et al. studies by using a more time-efficient format than the match-to-sample (MTS) procedure used by Carpentier et al. (this will be discussed at greater length further ahead). In the present study, two of the stages (Stages 2 and 4) addressed the relating of relations; Stage 4 demonstrated the RFT definition of analogy or deriving relations between relations, while Stage 2 tested the nonarbitrary precursor to analogical responding.

In Stage 4, the total score, as well as individual substage scores, show significant correlations with age, IQ score, and total assessment scores. The total score for arbitrary analogical responding showed a slightly stronger correlation with IQ performance than basic arbitrary relations (Stage 3), while basic arbitrary relations showed a slightly higher correlation with age compared to arbitrary analogical relations. These data provide further evidence that analogical relations are undoubtedly tied with intellectual potential.

The results for Stage 4 showed a marked difference in scores between the 4- to 5-year-old cohort and the 5- to 6-year-old cohort. These findings contribute to the extant RFT research suggesting a developmental divide in the acquisition of analogical ability at around five years of age (Carpentier et al., 2002, 2003). The 5-year-old participants in the Carpentier et al. studies demonstrated arbitrary equivalence-equivalence and non-equivalence-non-equivalence only after additional compound-matching training with the trained relations between elements (e.g., A1B1-A3B3). The mean scores for the 6- to 7-year-old cohort were better again and were more evenly distributed across frames. In light of the strong correlation between basic arbitrary relations and age mentioned above, it is possible that an additional year of practicing deriving relations (Stage 3) facilitated relating relations within relations in Stage 4.

Unlike the Carpentier et al. studies, however, the present study examined both nonarbitrary (Stage 2) and arbitrary (Stage 4) analogical responding across five frames, providing further insight into the developmental sequence and acquisition of analogical responding. Total relational assessment scores for Stage 2 (nonarbitrary analogical relations) showed strong, significant correlations with age, IQ score, and total assessment scores. Data were further analysed for patterns of responding by age group to identify at which age analogical responding is acquired. Stage 2 results showed a gradual improvement in scores by age and suggested the acquisition of analogy, at least in coordinate and difference relations, at five years of age. This is the first RFT study to focus on nonarbitrary analogy in addition to arbitrary analogy. In the domain of comparative psychology, more work has been done testing relating of nonarbitrary relations in nonhumans, and indeed, various species have been found to pass such tests. For example, chimpanzees (Gillan et al., 1981), crows (Smirnova et al., 2015), and baboons (Fagot & Maugard, 2013) have passed nonarbitrary analogical (i.e., relating relations) tasks. Additionally, cognitive-developmental researchers have also conducted such work (e.g., Christie & Gentner, 2014). The present study begins to bridge that gap in the RFT-based study of nonarbitrary analogy in young children. From an RFT perspective, of course, nonarbitrary analogy is not full analogy (i.e., deriving relations among arbitrary relations). However, it is an important repertoire and should be trained as a prerequisite skill before arbitrary analogy. The strong correlations between nonarbitrary and arbitrary analogy total scores in the present study provide evidence for this analysis, and future researchers could further evaluate the effects of training nonarbitrary analogy if it is a weak or missing skill in young children.

Finally, an additional point that might be made regarding the importance of analogy concerns the data from Participant S26, an outlier in the 3- to 4-year-old cohort. Participant S26's full-scale IQ score was 132; without his data, the mean IQ score for all participants was 112. Interestingly, Participant S26's data are similar to those of his cohort for the nonarbitrary relations and nonarbitrary analogy stages. However, his scores were above the mean in arbitrary relations and arbitrary analogy; he was the only participant under four years of age to respond correctly to arbitrary analogy tasks, thus positively skewing the results of the 3- to 4-year-old cohort. He also outperformed the 4- to 5-year-old participants in the arbitrary analogy stage. His data provide further evidence that arbitrary analogical responding is strongly tied with intellectual potential.

Relational Responding

There is now considerable proof that arbitrary relational responding is the principal behaviour characterising human language and cognition (Dymond & Roche, 2013; Fryling et al., 2020; Zettle et al., 2016). However, despite the evidence, research on the normative development of relational responding, including the sequence of relational frame acquisition, is as yet limited. The results in this study showed significant correlations between relational assessment scores and age, adding to the extant research showing that relational responding strengthens as a function of age in typically developing children (e.g., Dixon, Rowsey, et al., 2017; McHugh et al., 2004; Mulhern et al., 2017). In addition, the correlations observed in this study between relational performance and IQ score support and extend the growing body of research showing that relational framing supports linguistic and cognitive performance (e.g., Cassidy et al., 2011, 2016; Dixon et al., 2018; Gore et al., 2010; Hayes & Stewart, 2016; O'Hora et al., 2005, 2008; O'Toole & Barnes-Holmes, 2009; Moran et al., 2014; Mulhern et al., 2017). The strong correlations observed in this study indicate that the relational assessment employed here could be a reliable, relatively comprehensive tool for assessing relational skills in young children. Lastly, an important function of this study was to focus in particular on analogical reasoning, and the significant correlations found between analogy and IQ performance provide further evidence that analogy is a particularly relevant pattern of relational framing concerning intellectual potential (e.g., Sternberg, 1977).

Relational Responding and Age

The relational assessment developed in the present study examined five relational frames, including coordination, comparison, opposition, temporality, and hierarchy at both nonarbitrary and arbitrary levels. Performance on each frame was examined independently within and across stages in an effort to elucidate frame development across age. Results showed relatively predictable patterns wherein total relational assessment scores increased as a function of age, and average scores for each stage showed clear, stable improvements across age cohorts. A strong correlation was found between total assessment score and age, and further analyses showed significant correlations between total scores for each stage and age. These correlations suggest, consistent with RFT, that the capacity to engage in relational responding (both overall and across different frames) is established and strengthened via ongoing exposure to the typical socio-verbal environment.

The biggest age-based improvement in relational assessment scores between the youngest and oldest cohorts occurred in Stage 3 (arbitrary relations) and Stage 4 (arbitrary analogy). In these two stages, the gap in terms of overall correct responding between the 3- to 4-year-old and the 6- to 7-year-old cohorts was substantial, whereas the difference in scores was not as profound in Stages 1 and 2. These data suggest that there is already substantial development in nonarbitrary relations (as in Stage 1) before the age of three years, resulting in relatively less room for improvement after age three. Apart from finding a general improvement based on age, we also found particular discontinuities between specific age cohorts within stages. For example, in Stage 2 (nonarbitrary analogy), we noted a large gap in performance between the 4- to 5-year and 5- to 6-year-old children. These data may indicate that the transition from the 4- to 5-year-old to the 5- to 6-year-old age range results in a particular improvement in nonarbitrary analogy. In Stage 3 (arbitrary relations), the biggest age-based improvement in relational assessment scores occurred between the 3- to 4-year-old and 4- to 5-year-old cohorts, and there was almost no change between the 4- to 5-year-old and 5- to 6-year-old cohorts. These data might suggest considerable development in arbitrary relational framing between the 3- to 4-year-old and 4- to 5-year-old cohorts. This is a particularly important finding because the development of relational framing as a generalised operant across frames has not previously received much attention. One previous study (Mulhern et al., 2017) investigated the acquisition of hierarchical relations. Mulhern et al.

found that participants in the 3- to 4-year-old cohort did not show arbitrarily applicable responding in accordance with hierarchy, while participants in the 5- to 6-year-old cohort showed arbitrary hierarchical framing relatively robustly. These patterns of development across frames should be considered when designing an RFT-based curriculum focused on training relational framing, a point we discuss further below.

With respect to analogy (relating of relations) in the present study, data from Stage 4 (arbitrary analogy), similar to that from Stage 2 (nonarbitrary analogy), showed that the biggest age-based improvement in relational assessment scores occurred between the 4- to 5-year-old and 5- to 6-year-old cohorts, and there was almost no difference between the 3- to 4-year-old and 4- to 5-year-old cohorts. The fact that there appears to be a substantive improvement in analogical ability between the 4- to 5-year-old and 5- to 6-year-old participants at both the nonarbitrary and arbitrary levels supports earlier research (e.g., Carpentier et al., 2002, 2003), which also found an emergence of analogical reasoning around five years of age.

Within-stage analyses also showed some noteworthy correlations with respect to particular relational frames and age. For example, in Stage 1, the only significant correlation with age was the comparative relation; in Stage 2, coordination and hierarchy correlated with age; and, in Stages 3 and 4, all substages correlated with age. These analyses suggest some interesting findings to which we will return later.

Relational Responding and IQ

The second area of concentration in this study looked at the correlations between relational ability and IQ (see also Cassidy et al., 2011; Dixon et al., 2018; Gore et al., 2010; Moran et al., 2014; Mulhern et al., 2017; O'Hora et al., 2005, 2008; O'Toole & Barnes-Holmes, 2009). The present study represents an advance over these studies in terms of its scope; we examined derived relational responding across multiple frames and included typically developing children across different age cohorts.

The significant correlations observed in this study between relational performance and IQ support and extend the growing body of research finding that framing relationally is a requisite for linguistic and cognitive performance. In the current study, the correlations between IQ and responding in the arbitrary relational stages were more significant than those between IQ and responding in the nonarbitrary stages. Furthermore, within-stage analyses also showed significant correlations between all substage scores and IQ in arbitrary Stages 3 and 4. These correlations suggest that consistent with previous RFT research (e.g., Cassidy et al., 2011, 2016), there is a strong link between relational framing and intellectual potential. Additional within-stage analyses showed significant correlations between IQ and both nonarbitrary comparison and hierarchy scores in Stage 1. These findings are consistent with previous research noting the importance of comparative relational responding for intellectual potential (Cassidy et al., 2011, 2016). There were also significant correlations between Stage 2 coordination, comparison, and hierarchy scores, and IQ, findings consistent with the literature on analogy and intelligence (Christie & Gentner, 2014; Gentner & Christie, 2010; Goswami & Brown, 1989; Sternberg, 1977). Upon further examination, comparison and hierarchy were the only substages across all four stages that showed a consistent correlation with raw IQ as revealed by both SB5 total score as well as SB5 subscale scores. This pattern suggests that comparison and hierarchy may be particularly important foundational intellectual skills and should be considered for further investigation.

Normative Data

One point that we presented in the introduction was the fact that despite substantial evidence showing that relational framing is a core skill underlying cognitive ability, research on the normative development of relational responding and sequence of frame acquisition is as yet limited. In this study, we investigated the normative development of multiple relational frames in participants aged 3 to 7 years old. This is one of the first attempts to provide a cross-sectional, comprehensive look at the acquisition of frames in young children.

The relational assessment traced the acquisition of nonarbitrary to arbitrary relational responding and nonarbitrary to arbitrary analogical responding across frames of coordination, comparison, opposition, temporality, and hierarchy. In line with RFT, the normative data suggest that the ability to engage in relational responding (generally and across different frames) is acquired gradually throughout childhood (at least among typically developing children). The cross-sectional data suggest that typically developing children acquire the coordination relation first. This further supports previous suggestions that the coordination relation is the most common and ubiquitous pattern of relational responding and that its acquisition may facilitate the development of other relational patterns in a child's developing verbal repertoire (Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Hughes and Barnes-Holmes, 2016).

The current data also support the importance of comparative responding in accordance with age, IQ, and total relational assessment score. Nonarbitrary comparison was the only nonarbitrary relation that showed significant correlations with age, IQ score, total relational assessment score, and total scores for each stage. These findings suggest that nonarbitrary comparative relations may be an especially relevant relational pattern for intellectual potential, the emergence of arbitrary relational framing, and analogical responding. Furthermore, these data also suggest that the nonarbitrary comparative relation may play an important role in additional language development. Future researchers could further investigate training nonarbitrary comparative relations if they are weak or missing in a child's verbal repertoire in order to facilitate additional language development. The present data also suggest that arbitrary comparison is particularly important; for example, there were significant correlations between age, IQ, and total assessment score and comparison at every stage with the exception of Stage 2 comparison and age. Additionally, the comparison data show that arbitrary comparative relations had significant correlations with arbitrary coordination, opposition, temporality, and hierarchy; with total scores for each stage in the assessment; and, with all frames in Stage 4, arbitrary analogical responding. These findings provide further evidence that the comparative relation may be especially important for more complex language performance. This also echoes the finding from the Cassidy et al. (2011, 2016) studies that comparative relations play a key role in intellectual potential. Cassidy et al. found significant increases in mean IQ score after training multiple stimulus relations, including comparison. Finally, in both Stages 1 and 3 of the present study, the comparison substage was strongly correlated with all substages in Stage 4. These data suggest that nonarbitrary and arbitrary comparative relational repertoires are important intellectual skills and may be predictive of analogical ability. Future researchers may want to further investigate the effects of training nonarbitrary and arbitrary comparative relations in students both with and without language delays.

The present cross-sectional data also showed that scores were lowest for temporal relations across all age cohorts, suggesting that the temporal relation is acquired after the other tested frames. These findings are consistent with cognitive research, which has found that temporal understanding emerges relatively late in development (McCormack & Hoerl, 2017; Pyykkönen & Järvikivi, 2012). A noted advantage of the RFT approach adopted in the present study, however, is the analysis of temporal understanding in terms of functional units. The fact that we can conceptualise temporal understanding in terms of relational frames means that we can target this repertoire for training and assess for potential improvements on this basis. Within the RFT literature, temporal framing has previously received less empirical attention than other relations (Hughes & Barnes-Holmes, 2016). It shares the same basic pattern as comparative relations in that it entails responding to events in terms of their directional displacement along a specified dimension. However, time is inherently more abstract than other comparatives, such as size (Hayes, Gifford, et al., 2001), and requires discriminating successive changes in time. Thus, conceptualising the physical dimension along which temporal comparatives are arranged requires a more complex verbal repertoire and metaphorical understanding. This might explain why temporal framing appears to be acquired later than other frames from an RFT point of view. Given the importance of this repertoire, however, further research on testing and training of both nonarbitrary and arbitrary temporal relations is warranted.

In the present study, with the exception of opposite relational responding, participants performed better on the nonarbitrary stages than the arbitrary stages, thus supporting RFT and previous research (Berens & Hayes, 2007; Gorham et al., 2009). It is generally accepted that nonarbitrary responding is required for the acquisition of arbitrary relational responding (Rehfeldt & Barnes-Holmes, 2009). Previous research has found that participants who did not learn to derive a particular pattern of arbitrary relations were aided in doing so via additional training in the corresponding pattern of nonarbitrary relations (Berens & Hayes, 2007; Gorham et al., 2009). Similarly, in the PEAK assessment and curriculum model, Dixon et al. (2014) train simple and complex verbal relations using nonarbitrary stimuli first and then progress to equivalence and transformation training.

Given the foregoing, one perhaps surprising outcome of the present study was that participants scored higher in the arbitrary opposition stage than in the nonarbitrary opposition stage. We think it is possible that children scored poorly on nonarbitrary opposition because opposite relations are not typically systematically trained as two dichotomous stimuli on a specified physical dimension on a continuum (Hughes & Barnes-Holmes, 2016). Instead, opposite relations are more likely to be taught as an intraverbal response (e.g., the opposite of tall is short, big/small, full/empty), and thus, participants may be more familiar with the arbitrary format (i.e., what is the opposite of many/few?) than the nonarbitrary format as in the present study. However, it remains unclear whether participant performance was idiosyncratic to the test or idiosyncratic to their education. In any case, opposite relations are empirically important, as shown by Cassidy et al. (2011, 2016), for example, who found that training opposite and comparative relations led to improvement in IQ scores. Future researchers may want to investigate systematically training opposite relations as functional units, as defined by RFT, at both nonarbitrary and arbitrary levels.

Relational Assessment Format

One potentially important aspect of the present study was the assessment format. Previous RFT studies have used a number of different methodologies to examine and train relational framing in young children, including, for example, the matching-to-sample (MTS; see Rehfeldt & Barnes-Holmes, 2009) procedure and the relational evaluation procedure (REP; see Barnes-Holmes et al., 2001; Stewart et al., 2004). Matching-to-sample procedures require extensive baseline training prior to starting any testing or training. For example, in the Carpentier et al. (2002) study, the five-year-old participants required an average of 234 baseline trials in Experiment 1 before training and testing could begin. The REP, on the other hand, is a more efficient method in which multiple exemplars of relational patterns can be easily generated and readily presented without requiring any prerequisite match-to-sample training. This methodology allows participants to report on or evaluate sets of arbitrarily applicable relations defined by various sets of contextual cues. The testing format utilised in the current study was employed in the context of the multi-stage relational assessment that allowed testing of a range of different types of relations across different levels of complexity using the REP. The relational networks in the present study were displayed on the screen during each trial for participants to derive relations and relate relations among relations in the arbitrary stages, thus omitting the need for time-consuming MTS pre-training.

To make the present relational assessment equally accessible to all participants, including younger cohorts unable to read, a novel procedure was implemented, which excluded any textual instructions or tasks. Instead, this measure used simple, monochromic shapes and single, alphabetic letters to indicate the contextual cue (i.e., *S/D*, *M/L*, *S/O*, *B/A*, and *C/I* for same/different, more/less, opposite, before/after, and contains/inside, respectively). In addition, the letter identifying the contextual cue was paired with optional audio stimuli, for example, ‘same’ for S.

A functional analysis of intelligence as relational skills, such as the analysis provided by the relational assessment in this study, provides useful information on the relational skills that contribute to intelligent behaviour as assessed on IQ tests. Considering the predictive validity of IQ scores (Sternberg et al., 2001), understanding the variables that influence intelligent behaviour is a socially valid endeavour. Furthermore, the extant RFT research on relational training and intellectual performance demonstrates that relational framing can be trained and strengthened (e.g., Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Barnes-Holmes, Barnes-Holmes, Smeets, et al., 2004; Belisle et al., 2016; Berens & Hayes, 2007). Thus, identifying weaknesses in relational ability could provide the scaffolding needed for a relational curriculum (discussed in more detail below). The relational assessment protocol used in the present study can provide a useful instrument in such work going forward.

Despite its advantages, the relational assessment measure used herein may also need some further refinement. Analyses of the present data reveal certain interesting trends; nevertheless, we have to bear in mind that the assessment format may have made some relations more difficult than they would be if we had given participants a symmetry test or a match-to-sample task, for example. In addition, because the assessment is constrained in what it tests in each particular frame, we can make some comparisons between individual frames, but it is hard to be absolute or definitive about relational ability. For example, 1) we can clearly see that across the four stages, coordination is a simpler relation than the other relations; 2) in Stage 3, we can visibly see an improvement in substages across age; and 3) in Stage 4, we can see the acquisition of analogical responding at age five, but other comparisons are not as conclusive. However, the primary purpose of the assessment is to provide a general overview of a child's nonarbitrary and arbitrary relational repertoire. To get a more comprehensive measure of a child's relational abilities, future researchers could develop a relational protocol that systematically assesses direct relations, entailed relations, and transformation of stimulus function with more scope and depth. For example, this could be done by adding more detailed tasks to each substage, one stage at a time, or by focusing specifically on the evolution of nonarbitrary to arbitrary relational framing across age groups, one frame at a time. Furthermore, additional relations such as distinction and deixis could be included in future relational assessments. Approaching relational training in such a functionally specified way will enable language protocols to assess and train more generative, flexible, and complex language.

A second potential limitation of the relational assessment may be the omission of distinction as an independent frame. In the present study, distinction was not tested in Stage 1; and in Stages 2-4 distinction was included in the coordination substage to provide a comparison option. More work on distinction is warranted, and future researchers may consider testing coordination and distinction more systematically and providing independent analyses for both frames.

A third potential limitation of the relational assessment may be the linear versus nonlinear and same versus mixed presentation of the arbitrary relational networks in Stages 3 and 4. In a linear-same series, stimuli are presented sequentially (e.g., A – B – C), and the contextual cue between relata is the same. For example, in a linear-same comparative relational network, A is more than B, and B is more than C, or A > B, B > C (Fienup &

Brodsky, 2020; Vitale et al., 2008). In a linear-mixed series, stimuli are presented sequentially, but the contextual cue is not the same between relata, for example, A > B, B < C. In nonlinear series, stimuli are not presented linearly (e.g., A – B, C – A); in a nonlinear-same series, the contextual cues are the same between relata, and in a nonlinear-mixed series, the cues are not the same. For example, A > B, C > A and B > A, C < A, respectively (Vitale et al., 2008). In Stage 3 in the present study, the coordination substage included linear-same, linear-mixed, and nonlinear-same relational networks, and all other substages (comparison, opposition, temporality, and hierarchy) included only linear-same relational networks. To distinguish between sameness and difference, the difference relation had to be included in the coordination relational networks, thus, resulting in mixed networks. For all substages, including coordination, however, questions regarding relations in the network always progressed from directly trained, to mutually entailed, to combinatorially entailed relations.

In Stage 4, the coordination and opposition substages included linear-mixed networks, and the comparative, temporal, and hierarchical substages included nonlinear-mixed networks. Deriving relations from nonlinear networks is more difficult than deriving from linear networks (Hunter, 1957; Vitale et al., 2008). However, all substages included mixed contextual cues, and as in Stage 3, trials progressed from directly trained, to mutually entailed, to combinatorially entailed analogical relations. Future replications of Stages 3 and 4 should include systematic transitions from linear-same-mixed to nonlinear-same-mixed relational networks in order to gain more detailed information regarding participants' relational skills.

Future Curriculum Development

Considering the potential for improving language and cognition, further research into relational assessment and training is clearly warranted. Up until recently, most empirical work on verbal behaviour has largely been influenced by Skinner's (1957) analysis of verbal behaviour (Dymond et al., 2010; Dixon, Belisle, et al., 2017). Consequently, commonly used language assessments, such as the Assessment of Basic Language and Learning Skills – Revised (ABLLS-R; Partington, 2008) and the Verbal Behaviour Milestones and Placement Program (VB-MAPP; Sundberg, 2008), are based on Skinner's analysis of verbal behaviour; thus, training focuses on the basic verbal operants (i.e., echoics, mands, tacts, intraverbals) with little attention to more complex verbal behaviour. More recently, Dixon et al. (2014, 2018) and Dixon, Rowsey, et al. (2017) provided experimental work assessing and training more complex language. The present study contributes to the work on relational language assessment.

Previous relational training studies (e.g., Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Belisle et al., 2016; Berens & Hayes, 2007) showed that derived relational responding could be brought under operant control. A functional analysis of young or developmentally delayed children's existing relational abilities would provide the framework for a robust, flexible, and individualised RFT-based curriculum. The present study is one of the first studies to look at the sequence of acquisition of multiple frames in young children; these findings could be a valuable reference in curriculum design.

Related to this point, one possible limitation in the current study is the representativeness of the sample population. All participants recruited for this study attended the same private school in New Jersey. As mentioned, the average IQ score for participants in this study without the outlier was 112 (113 with the outlier participant), which is slightly above the average range of 90-109 (Roid, 2003). Future research using the present protocol to assess relational framing in young children should include substantially larger numbers both overall as well as within each of the age cohorts. It should also recruit from a wider variety of educational institutions such as public schools and disadvantaged schools. Such extensions of the present work would arguably constitute more representative testing and provide more generalisable results.

Despite this point, the present data are informative in terms of the development of a possible RFT-based curriculum. For example, the present data show correlations between age and both nonarbitrary and arbitrary relational ability. These normative data could help inform a developmentally sequenced relational curriculum; for example, a logical training sequence would include training increasingly difficult or complex levels of nonarbitrary relational frames before training arbitrary relational frames. Another area of warranted research includes testing and training the transition from nonarbitrary to arbitrary relations; a comprehensive relational curriculum should include training-for-transition procedures.

Additional within-stage analyses show some interesting results regarding particular relational frames that could be taken into account when designing relational programmes. For example, in Stage 1, there was a significant correlation between comparison and age, IQ, total relational assessment scores, and total substage scores for each substage. Future researchers could investigate training nonarbitrary comparative relations in order to facilitate the development of other frames. In Stage 2, nonarbitrary analogical frames of coordination and hierarchy correlated with age; therefore, it may be worth training these frames first in nonarbitrary analogy programmes. In Stages 3 and 4, all substages correlated with age, and more specifically, the data showed at which age particular frames were acquired. In Stage 3, the biggest age-based improvement in arbitrary relational scores occurred between the 3- to 4-year-old and 4- to 5-year-old cohorts. These data are consistent with the Mulhern et al. (2017) study, which found that 3- to 4-year-old participants did not respond correctly on tests of hierarchical framing, whereas participants aged five and older performed better on these tests. Based on these data, considerable development in arbitrary relational framing occurs between ages four and five. Additionally, the data from Stage 3 in the present study indicate that coordination is acquired first, and temporality is acquired last. Thus, a relational language programme could introduce relational training around the developmental age of four and start with training the relational frame of coordination first. The previously mentioned VB-MAPP (Sundberg, 2008) is based on the verbal repertoires of typically developing children up to four years old; thus, the development of a robust relational assessment and curriculum suitable for the developmental age of four-plus is warranted.

Another consideration for inclusion in a relational curriculum is analogical responding. Considering the significant correlation between analogical responding and IQ, identifying and training deficits in analogical reasoning as a relational repertoire has the potential to further strengthen relational programmes. In Stage 4, analogical responding developed around the age of five; thus, training arbitrary analogy should begin around the developmental age of five. Again, the frame of coordination should be introduced first.

Relational frame theory proposes that fluent and flexible derived relational responding may underlie much of human cognition (Hayes et al., 2001); thus, a functional analysis of existing relational abilities would facilitate designing effective relational programmes. Furthermore, several studies have investigated the effects of training relational responding and have found that training in relatively few frames results in improvements across a wide range of intellectual and language processes. For example, Cassidy et al. (2011, 2016) and Hayes and Stewart (2016) found significant increases in IQ scores and intellectual performance for all participants after completing automated relational training. Mulhern et al. (2018) trained hierarchical framing in young children and also saw increased scores in assessments of language and categorisation. Therefore, by testing a child’s relational abilities and having a better understanding of the normative development of relational frames, we can design a curriculum that targets specific relational frames and build a functional relational repertoire without omitting the necessary component relational skills. Furthermore, by providing a general overview of a child’s relational repertoire, the assessment in this study could be the basis for an RFT-based curriculum for children who present with learning difficulties or with developmental and intellectual disabilities.

Conclusion

Analogical reasoning is frequently applied as a metric of intelligence (Sternberg, 1977), and as a measure to predict academic success. Thus, a functional analytic evaluation of analogy as relations among relations may provide a critical, heretofore missing, component of language assessment. Approaching analogical reasoning in this functionally specified way can facilitate practical intervention for deficiencies in analogical responding, such as in young children or in children with developmental delays. Future studies should examine RFT-based training protocols targeting analogical relations found to be missing or weak. The results of the present study add to extant literature that derived relational responding and intelligence are related, providing further support for a functional analysis of intelligent behavior embedded in a derived relational account of human language and cognition. There is a growing body of literature showing that when relational skills repertoires are improved, IQ scores also increase, thus interventions which target relational repertoires should be a training priority. Thus, the successful assessment and training of analogical relations and its effect on IQ might may be even more advantageous than currently touted.

Chapter 4: Study 2. Testing and Training Analogy in Young Children

Introduction

As mentioned in the previous chapter, 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 peach is to pineapple as goat is to horse. In this case, peach and pineapple participate in an equivalence relation in the context of fruit; and goat and horse participate in an equivalence relation in the context of animal, and thus, because these are both equivalence relations, we can derive a relation of equivalence between the relations themselves.

To model this phenomenon, Barnes et al. first trained and tested four three-member equivalence relations in adults and 9-year-old children. A matching-to-sample (MTS) procedure was used to train conditional discriminations amongst three-letter nonsense syllables (coded using alphanumeric designations) as follows: A1→B1, A1→C1, A2→B2, A2→C2, A3→B3, A3→C3, A4→B4, A4→C4. Next, participants were tested for the derivation of the following four untrained combinatorially entailed relations: B1↔C1, B2↔C2, B3↔C3, B4↔C4. After participants passed equivalence tests, they were tested for the derivation of equivalence relations between equivalence relations themselves, which the researchers referred to as equivalence-equivalence responding. Compound stimuli comprised either two nonsense syllables that were equivalent or two that were non-equivalent. Participants were required to choose an equivalent pair in the presence of an equivalent pair (i.e., equivalence-equivalence) and a non-equivalent pair in the presence of a non-equivalent pair (i.e., non-equivalence-non-equivalence). For example, given B3C3 and B3C4 as comparisons, if the sample was B1C1, then they had to choose B3C3, whereas if B1C2 was the sample, they had to choose B3C4. All participants (i.e., both the adults and the 9-year-old participants) successfully related relations, and thus this constituted a basic analogy model.

Several studies since Barnes et al. (1997) have extended this model of analogy in several respects (Barnes-Holmes et al., 2005; Carpentier et al., 2002, 2003, 2004; Ruiz & Luciano, 2015; Stewart et al., 2001). One strand of research of particular interest for the present purpose focused on analogical responding in young children (e.g., Carpentier et al., 2002, 2003). Carpentier et al. (2002) found that adults and 9-year-old participants readily showed equivalence-equivalence (i.e., as in the original study), but 5-year-old children, while readily passing equivalence, initially failed to show equivalence-equivalence without additional prompting. More specifically, the 5-year-old children required pre-testing with compound-compound-matching tasks involving trained (as opposed to derived) relations (e.g., A1B1-A3B3 and A1B2-A1B3) before they successfully passed 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. Despite receiving considerable additional training, only two of the 18 5-year-old participants successfully passed this task.

The present study sought to extend Carpentier et al. (2002) and further examine the acquisition of analogical responding operationalised as derived relations between relations in 5-year-old children. It extended the earlier work first by attempting to directly train the skill of deriving relations between relations using a controlled multiple baseline design to demonstrate experimental control. In addition, in the present study, we used an alternative, more efficient format to assess and train the derived relational pattern than the match-to-sample (MTS) procedure used by Carpentier et al. Although often used in studies of derived relations, MTS procedures require extensive baseline training before any testing or training 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 training and testing could begin.

The RFT-based relational evaluation procedure (REP; see Barnes-Holmes et al., 2001; Stewart et al., 2004) offers a potential alternative to MTS. In the REP, participants are required to evaluate relational networks based on the presentation of contextual cues juxtaposed with relevant stimuli. Using this protocol, multiple exemplars of relational tasks can be readily generated and presented without requiring lengthy prerequisite training. The REP has been successfully employed in several recent RFT-based studies to train relational framing in young children. For example, in the previously mentioned Cassidy et al. (2011) study, an automated AADRR assessment and training programme was based on the REP format (see also Cassidy et al., 2016; Hayes & Stewart, 2016). The automated programme 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?’). The successful use of this protocol by Cassidy et al. to efficiently train key patterns of relational framing to 8- to 12-year-old children, consequently boosting their intellectual performance, suggests the potential utility of the REP format in training relational framing in children, and hence we adopted this protocol in the current study.

In Study 1 of this thesis, we developed and tested a type of REP specially created to allow for assessing relational framing in very young children. In Study 1, we designed a comprehensive relational assessment to test various relational frames across four levels of responding, including nonarbitrary relations, nonarbitrary analogical relations, arbitrary relations, and arbitrary analogical relations. Assessment of the latter two stages of responding (i.e., arbitrary relations and analogical relations) was also required in the present study. To test these stages in Study 1, we taught children to respond to relational networks composed of monochromatic shapes (instead of nonsense words) juxtaposed with single letters as the relevant contextual cues (e.g., S for sameness, D for difference). For example, children were taught that given a green circle and a red circle separated by the contextual cue 'S', they should subsequently treat the green and red circles as equivalent. In the analogical relations stage, the analogical stimuli were compound elements composed of monochromatic shapes presented below the relational network. This format allowed young children, not yet able to read, to report on and evaluate multiple exemplars of arbitrarily applicable relational networks defined by specifically selected contextual cues. This methodology was effective in assessing arbitrary relational responding as well as the relating of arbitrary relations in these children.

The aim of the present study was to extend the work of Carpentier et al. (2002) as well as that of Study 1 by not just assessing but also training the relating of relations in a multiple baseline design with 5-year-old children using an adaptation of the REP type protocol used in Study 1.

Method

Participants and Setting

Participants were three typically developing children enrolled in a private elementary school on the east coast of the United States. P1 was a male aged 5 years and 4 months, P2 was a female aged 5 years and 11 months, and P3 was a female aged 5 years and 9 months. Participants were selected for inclusion based on their performance on an adaptation of the relational assessment used in Study 1 (see Figure 4.1 for a summary version of the pre-assessment). The researcher administered all probe and training sessions in an otherwise unoccupied classroom of the participants' school during school hours. The researcher sat next to the participant at a standard school desk. A second independent observer sat approximately three feet away from the desk on the other side of the participant with a full view of the participant and the computer screen.

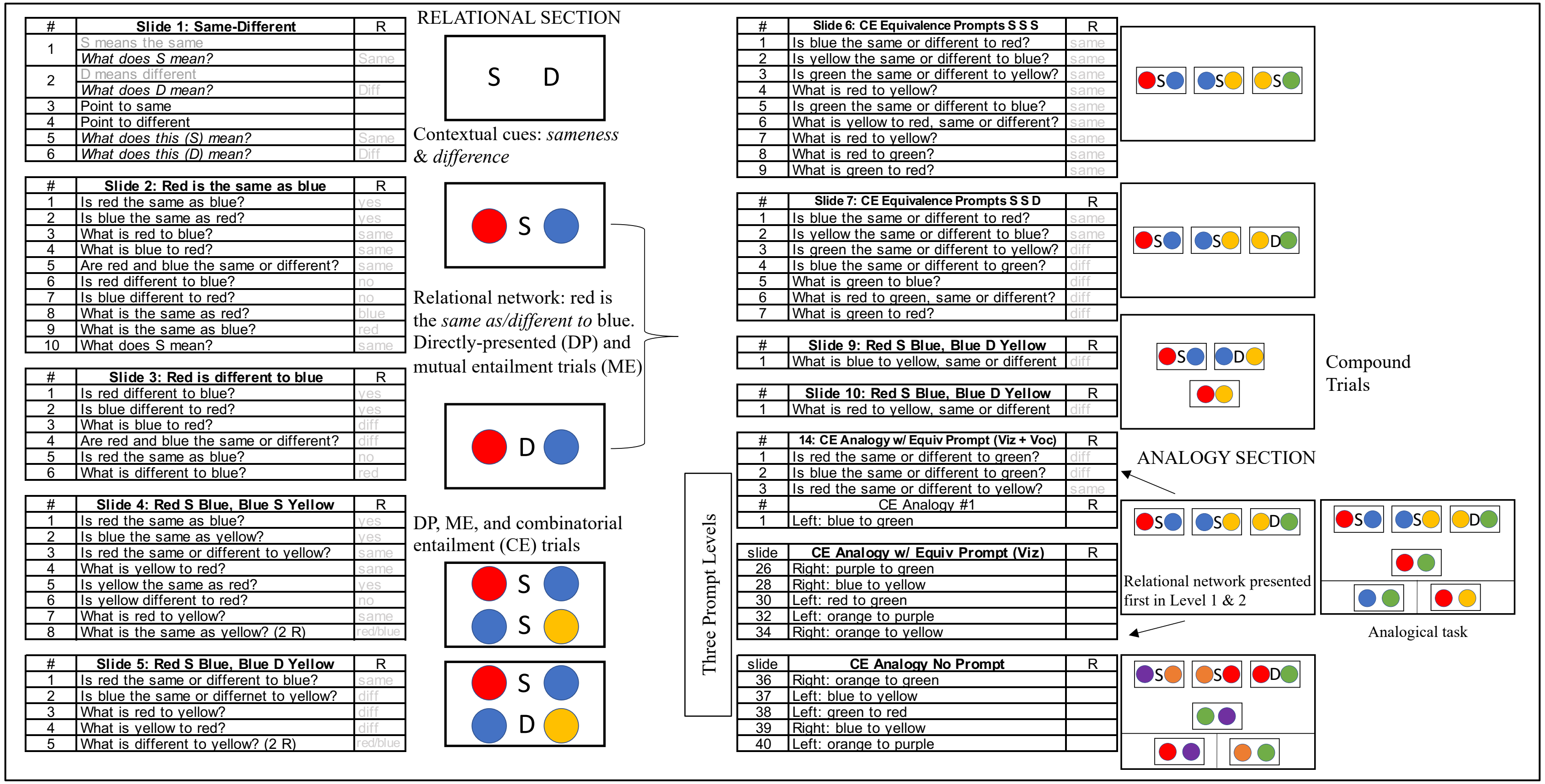
Ethical approval for the recruitment of participants was obtained from the research ethics committee of the lead researcher's host institution. Consent for conducting the study was also obtained from the principal of the school. Caregiver consent was obtained for each child who participated, and verbal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

Experimental Design

A combination of a multiple baseline design across participants and a multiple probe design across behaviours was used in this study. The study included a relational pre-assessment for screening potential participants; a baseline condition in which analogy was tested; and a training condition in which analogical responding was trained, and generalisation probe trials were presented. For P2 to enter the training condition, the participant previously exposed to training (i.e., P1) had to meet the training criterion (i.e., scoring 100% correct twice consecutively on the trained repertoire) while for P3 to enter the training condition, P2 had to meet the training criterion.

Figure 4.1

Pre-Assessment: Sample of Relational and Analogical Sections



RELATIONAL SECTION

Contextual cues: *sameness & difference*

Relational network: red is the *same as/different to* blue. Directly-presented (DP) and mutual entailment trials (ME)

DP, ME, and combinatorial entailment (CE) trials

Three Prompt Levels

Compound Trials

ANALOGY SECTION

Relational network presented first in Level 1 & 2

Analogical task

#	Slide 1: Same-Different	R
1	S means the same	
	What does S mean?	Same
2	D means different	
	What does D mean?	Diff
3	Point to same	
4	Point to different	
5	*What does this (S) mean?*	Same
6	*What does this (D) mean?*	Diff

#	Slide 2: Red is the same as blue	R
1	Is red the same as blue?	yes
2	Is blue the same as red?	yes
3	What is red to blue?	same
4	What is blue to red?	same
5	Are red and blue the same or different?	same
6	Is red different to blue?	no
7	Is blue different to red?	no
8	What is the same as red?	blue
9	What is the same as blue?	red
10	What does S mean?	same

#	Slide 3: Red is different to blue	R
1	Is red different to blue?	yes
2	Is blue different to red?	yes
3	What is blue to red?	diff
4	Are red and blue the same or different?	diff
5	Is red the same as blue?	no
6	What is different to blue?	red

#	Slide 4: Red S Blue, Blue S Yellow	R
1	Is red the same as blue?	yes
2	Is blue the same as yellow?	yes
3	Is red the same or different to yellow?	same
4	What is yellow to red?	same
5	Is yellow the same as red?	yes
6	Is yellow different to red?	no
7	What is red to yellow?	same
8	What is the same as yellow? (2 R)	red/blue

#	Slide 5: Red S Blue, Blue D Yellow	R
1	Is red the same or different to blue?	same
2	Is blue the same or differnet to yellow?	diff
3	What is red to yellow?	diff
4	What is yellow to red?	diff
5	What is different to yellow? (2 R)	red/blue

#	Slide 6: CE Equivalence Prompts S S S	R
1	Is blue the same or different to red?	same
2	Is yellow the same or different to blue?	same
3	Is green the same or different to yellow?	same
4	What is red to yellow?	same
5	Is green the same or different to blue?	same
6	What is yellow to red, same or different?	same
7	What is red to yellow?	same
8	What is red to green?	same
9	What is green to red?	same

#	Slide 7: CE Equivalence Prompts S S D	R
1	Is blue the same or different to red?	same
2	Is yellow the same or different to blue?	same
3	Is green the same or different to yellow?	diff
4	Is blue the same or different to green?	diff
5	What is green to blue?	diff
6	What is red to green, same or different?	diff
7	What is green to red?	diff

#	Slide 9: Red S Blue, Blue D Yellow	R
1	What is blue to yellow, same or different	diff

#	Slide 10: Red S Blue, Blue D Yellow	R
1	What is red to yellow, same or different	diff

#	14: CE Analogy w/ Equiv Prompt (Viz + Voc)	R
1	Is red the same or different to green?	diff
2	Is blue the same or different to green?	diff
3	Is red the same or different to yellow?	same
#	CE Analogy #1	R
1	Left: blue to green	

slide	CE Analogy w/ Equiv Prompt (Viz)	R
26	Right: purple to green	
28	Right: blue to yellow	
30	Left: red to green	
32	Left: orange to purple	
34	Right: orange to yellow	

slide	CE Analogy No Prompt	R
36	Right: orange to green	
37	Left: blue to yellow	
38	Left: green to red	
39	Right: blue to yellow	
40	Left: orange to purple	

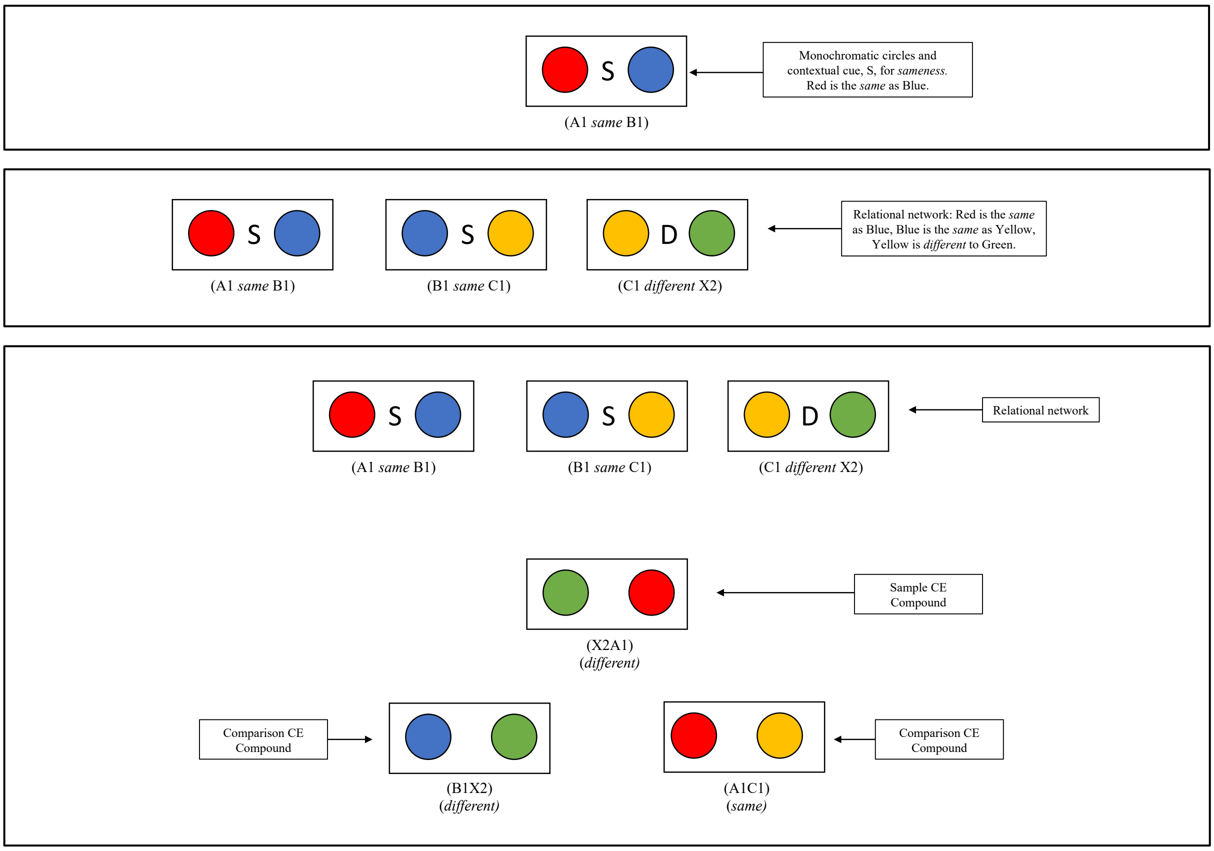
Note. In the CE Analogy w/ Equiv. Prompt (Viz & Voc) trials, the relational network was first presented without the analogical stimuli plus three CE relational trials followed by analogical stimuli; in the CE Analogy w/ Equiv. Prompt (Viz) trials, the relational network was first presented without the analogical stimuli followed by the analogical stimuli; in the CE Analogy No Prompt trials, the relational network and the analogical stimuli were presented on the screen simultaneously. Adapted Study 1.

Materials & Apparatus

A 13” MacBook running Microsoft PowerPoint was used to present trials. Individual stimuli included coloured circles (approximately 1 inch in diameter) and the letters ‘S’ for sameness and ‘D’ for difference relations in Calibri or Arial, size 40 font (see first panel in Figure 4.2). During the relational section of the pre-assessment, trials included an array of such stimuli at the centre of the screen that were designated by the experimenter as participating in a relational network. During the analogical section of the pre-assessment, as well as during the study probe and training sessions, similar to the relational section, trials included an array of stimuli that were designated as participating in a relational network. In this section, however, these appeared in the top portion of the screen while in the bottom portion of the screen there appeared a sample compound element below the relational network; and two comparison compound elements below the sample on the bottom left and right of the screen, separated by a black line (see third panel in Figure 4.2 for an illustrative example of the relational network and compound elements). The relational network in the top portion of the screen included six monochromatic circles and the relational cues, S for same and D for different, to delineate the relations between the circles (see second panel in Figure 4.2). For example, one possible array might be represented as follows: [Red Circle] [S] [Blue Circle], [Blue Circle] [S] [Yellow Circle], [Yellow Circle] [D] [Green Circle], which indicates that the red circle is the same as the blue circle, the blue circle is the same as the yellow circle, and the yellow circle is different from the green circle. The compound elements that appeared lower down the screen 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 third panel in Figure 4.2). For example, one such compound might be designated as [Red Circle][Yellow Circle]. Each slide had only one relational network and one set of analogical task stimuli.

Figure 4.2

Stimuli Format in Pre-Assessment and Multiple Baseline Conditions



Procedure

Overview

The following will provide procedural details of the relational pre-assessment and the multiple baseline component of the study. The pre-assessment was administered first in one session and took between 5 and 25 minutes to complete. Following the pre-assessment, participants entered the multiple baseline, including baseline sessions followed by training and multiple probe sessions. Sessions lasted between 10 and 40 minutes, and on average, included seven probe and training sets; a probe set took approximately two to four minutes to complete, and a training set took approximately one to three minutes to complete. All participants started baseline sessions at the same time and all participants took approximately five to six weeks to complete the study. Maintenance probe sessions were administered one month after each participant's last probe in the training condition was administered.

Pre-Assessment

The relational pre-assessment adapted tasks from the arbitrary relational and analogical stages in Study 1 of the present thesis. The pre-assessment included two main sections; the first section tested for derived relational responding, including mutual entailment and combinatorial entailment across five subsections increasing in difficulty. The second section tested for derived relations between combinatorially entailed relations (i.e., analogical responding) across three levels of prompting (refer to Figure 4.1 for a sample of trials and stimuli found in the pre-assessment). There were a total of 89 trials in the pre-assessment.

Section 1: Relational Tasks. The first six trials in the pre-assessment introduced the participant to the contextual cues themselves (i.e., S and D; refer to Figure 4.1). 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]. 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 16 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 (which we will hereafter refer to as mutual entailment) (e.g., ‘Is Blue the same as Red?’).

The next set of 29 trials included more than two stimuli and questions became increasingly difficult and required responding not only to ‘directly presented’ (DP) and ‘mutual entailment’ (ME) type questions but also to questions that required the combination of directly presented relations (hereafter referred to as combinatorial entailment; CE). The first set of questions in this section referred to a relational network in which only two sameness relations were presented; the second set referred to a relational network including both a sameness and difference cue; the third set included three sameness relations, and the fourth set included the relational network format that would be used in the study training sets involving the sequence: A is the same as B, B is the same as C, and C is different to D. An example of the latter set might be as follows: The relational network [Red Circle] [S] [Blue Circle], [Blue Circle] [D] [Yellow Circle] is presented on the screen, and then questions regarding combinatorially entailed relations (e.g., ‘Is Red the same/different to Yellow?’) are presented.

The next set of questions involved compound elements as described in the previous section, a stimulus composed of two side-by-side monochromatic circles from the relational network without the relational cues, S or D. For example, one such compound might be designated as [Red Circle][Yellow Circle]. In each trial, a relational network was presented at the top of the screen. Below the network, a white rectangle with a black outline contained the compound element (i.e., two differently coloured circles identical to two of the circles in the relational network). The researcher and the participants read the relational network together, and then the researcher said, ‘Look here (points to compound element), are they the same or different? Remember to look here (points to relational network) to help you figure it out.’ Participants had to refer to the network in order to correctly identify the relation between the stimuli (the circles) in the compound element. For example, the relational network [Red Circle] [S] [Blue Circle], [Blue Circle] [D] [Yellow Circle] is presented at the top of the screen, and the compound stimulus (e.g., [Red Circle][Yellow Circle]) is presented below the network; thus, the participant might derive that since red is the same as blue, and blue is different to yellow, therefore, red is different to yellow (the compound element). Each slide had only one relational network and one set of analogical stimuli.

To proceed to the analogical questions in the pre-assessment, the participant had to pass the last 14 combinatorial entailment and compound identification trials in the relational section of the pre-assessment at 80% correct. Only the last 14 trials of the relational section were included because the previous sections introduced the participant to the format and did not exclusively ask combinatorially entailed type or compound-type questions.

Section 2: Analogical Tasks. In the analogical section of the pre-assessment, the analogical trial stimuli (i.e., both the sample and two comparisons) included a relational network and three compound elements composed of two circles (see third panel in Figure 4.2). For example, a presented relational network might be: [Red Circle] [S] [Blue Circle]; [Blue Circle] [S] [Yellow Circle]; [Yellow Circle] [D] [Green Circle]; and read as, ‘Red is the same as blue, blue is the same as yellow, and yellow is different from green.’ The sample compound element presented below the relational network might be [Red : Green], and the comparison compound elements below the sample might be [Red : Yellow] and [Blue : Green]. On each trial, the researcher read the relational network to/with the participant and then 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?' In the example just given, the participant could look at the sample compound [Red : Green] and derive (i.e., based on the network) that since red is the same as blue, and blue is the same as yellow, and yellow is different from green, then red is different from green. They could then look at the comparison compounds [Red : Yellow] and [Blue : Green] and derive that since red is the same as blue and blue is the same as yellow, then red is the same as yellow (first comparison), and since red is the same as blue, and blue is the same as yellow, and yellow is different from green, then blue is different from green (second comparison). The correct (analogical) choice would be to select the second comparison (i.e., [Blue : Green]) because, as in the case of the sample stimulus, this involves a difference relation.

For all trials in the pre-assessment and in the multiple baseline conditions, the assessor first read the relational network to/with the participant and then asked the trial questions. If the participant requested or initiated independent reading when presented with the relational network, the researcher did not read with the participant.

The analogical section of the pre-assessment included three levels of prompting, which together represented a 'most-to-least' type prompting strategy. Each level included five combinatorially entailed (CE) analogy trials. Level 1 included CE analogies with visual and vocal equivalence prompts; Level 2 included CE analogies with visual prompts; and Level 3 included CE analogies with no prompts. Each level included five analogical trials. The procedure described in the previous paragraph was used in Level 3 (CE analogy with no prompts) and also in the training condition. In Levels 1 and 2, the relational network was first presented on the screen without the compound elements (i.e., the sample and the comparisons), and the researcher and participant read the relational network together. In Level 1, after reading the relational network, the researcher asked three combinatorial entailment questions that would refer to the relations of the compound elements in the ensuing analogy task. This prompt was implemented to further assess combinatorial entailment and support responding in the ensuing analogical task. After responding to the three CE questions, the analogical stimuli were presented below the relational network. The participant was given the instruction to look at the sample first and then to choose which comparison was like the sample. Thus, in Level 1, each trial included three CE questions in addition to the analogy trial. Feedback was provided for the CE questions; correct responses were praised, and incorrect responses were re-presented until the participant emitted a correct response. Feedback was not provided for the analogical questions. In Level 2, the researcher and participant read the relational network together first (there were no combinatorial entailment questions), and then the analogical compound stimuli were presented below the relational network as in Level 1. Level 3 was presented after Level 2. All levels were presented regardless of performance to confirm that the potential participant could derive combinatorially entailed relations but could not complete analogical tasks (i.e., relate relations).

The passing criterion was 100% (5/5) correct analogy trials at each level. Hence, the passing criterion for the full analogy section of the pre-assessment was 15/15 analogy trials. With regard to inclusion as a whole, potential participants were required to pass the last set of relational trials in the pre-assessment (14 trials) at 80% or more and to fail the analogy section (see Table 4.1 for a schematic presentation of the procedure).

Table 4.1

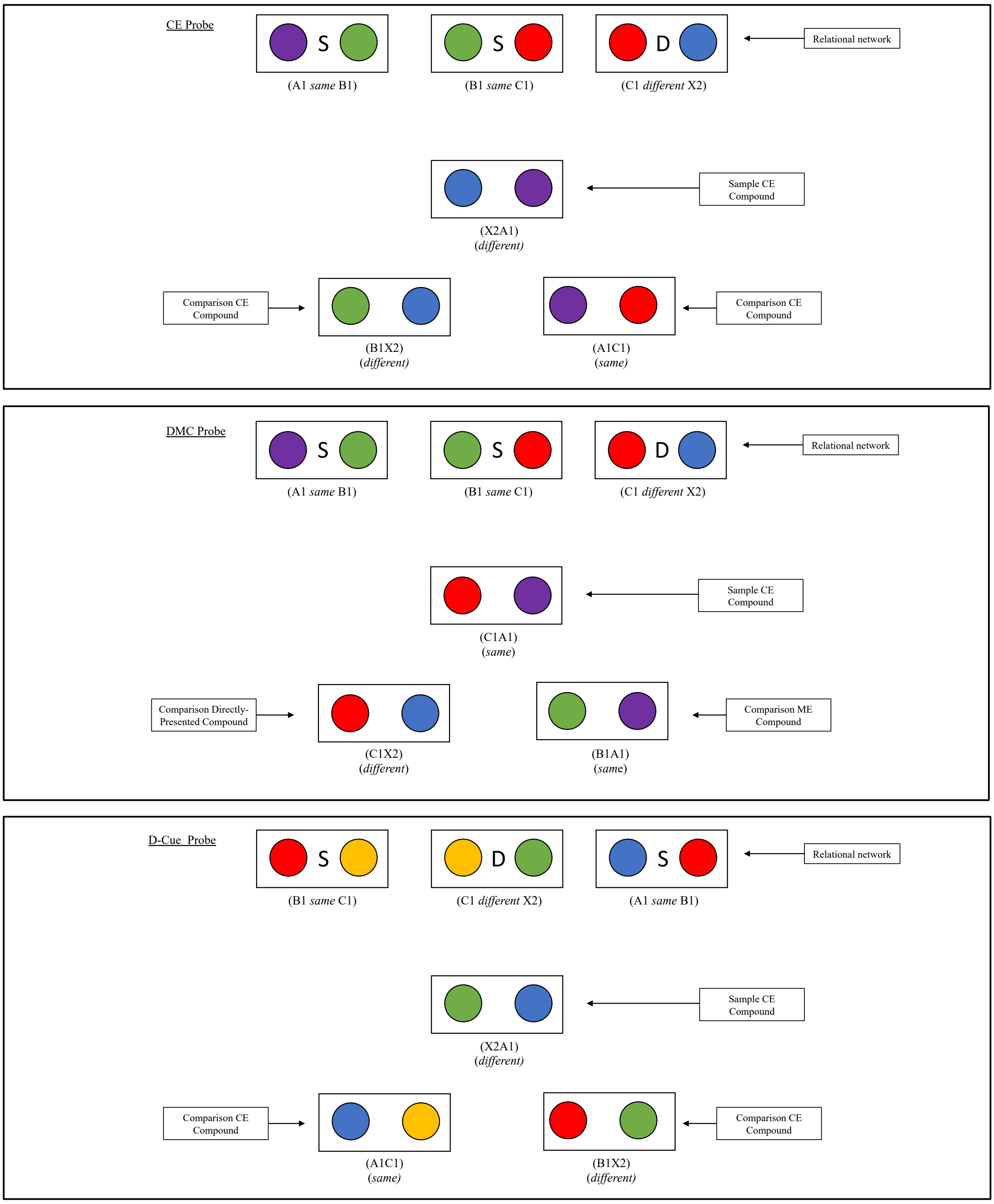
Schematic Presentation of the Pre-Assessment and Multiple Baseline Conditions

	Phase	Relational Targets	Passing Criteria %
Pre-Assessment			
	1.	Identifying contextual cues	100
	2.	Reading relational networks	100
	3.	Testing DP and ME relations	100
Inclusion Criteria	4.	Testing CE relations	80
	5.	Testing compound elements	80
		Analogy Probes	
	6.	Level 1: 3 CE trials + 5 analogy trials	100
	7.	Level 2: 5 analogy trials	100
	8.	Level 3: 5 analogy trials	100
Multiple Baseline Study			
		Baseline: CE Probe Set 1	100
Training & Probes		Training Phase 1: 3 CE trials + 1 analogy trial	100
		CE Probe Set 1: 5 trials	100
		Training Phase 2: 3 CE trials + 3 analogy trials	100
		Novel CE Probe: 5 trials	100
		DMC Probe: 5 trials	100
One-Month Maintenance Probes		Novel CE Probe: 5 trials	100
		D-Cue Probe: 5 trials	100
		Novel CE Probe	100
		Novel DMC Probe	100
		Novel D-Cue Probe	100

Note. DP = directly presented relations; ME = mutually entailed relations; CE = combinatorially entailed relations.

Figure 4.3

Three Analogy Probe Types: CE Probes, DMC Probes, and D-Cue Probes



Note. First panel: combinatorially entailed relations; second panel: mixed directly presented, mutually entailed, and combinatorially entailed relations; third panel: the difference cue changes location in the relational network plus combinatorially entailed relations.

Multiple Baseline Conditions

The multiple baseline components included a baseline condition including unreinforced probe sessions and a training/testing condition in which training and probe sessions were administered. The study included three different types of probes, including Combinatorially Entailed Analogy Probes (CE Probes); Directly Presented/Mutually Entailed/Combinatorially Entailed Analogy Probes (DMC Probes); and Difference-Cue-Shifted Analogy Probes (D-Cue Probes) (see Figure 4.3 for an illustrative example of each probe type).

In CE Probe sets, the relational network always included the presented relations A is the same as B, B is the same as C, and C is different to D. Combinatorial entailment was involved in deriving the relation between the two stimuli (i.e., two monochromatic circles) in all sample and comparison compound elements. During baseline sessions, the same CE Probe Set was administered for all sessions. Novel CE Probe sets were included for generalisation testing (see first panel in Figure 4.3).

The relational network in the DMC Probes followed the same pattern as in the CE Probe (i.e., [A [S] B], [B [S] C], [C [D] D]). However, the relations between the stimuli in the sample and comparison compound elements included a mixture (within as well as across trials) of directly related, mutually entailed, and/or combinatorially entailed relations. In directly related compounds, the relation between the stimuli was directly presented in the network and thus did not require any derivation (e.g., [AB] or [BC]). For example, a DMC Probe trial might include a sample with directly related stimuli, one comparison involving a mutually entailed relation, and a second comparison involving a combinatorially entailed relation (see second panel in Figure 4.3).

The critical distinction of the D-Cue Probes was that the format of the relational network did not follow the [A [S] B], [B [S] C], [C [D] D] pattern. Instead, the location of the difference cue varied on every trial in a probe set. There were always two sameness cues and one difference cue in the relational networks, but across D-Cue Probe trials, the difference cue could be the first, second, or third relational cue. For example, the difference cue is in the first position in the network A is different to B, A is the same as C, C is the same as D ([A] [D] [B], [A] [S] [C], [C] [S] [D]), while it is in the second position in the network A is the same as B, B is different to C, A is the same as D ([A] [S] [B], [B] [D] [C], [A] [S] [D]). As in the CE Probes, combinatorial entailment was required to derive the relation between the two stimuli in all sample and comparison compound elements (see third panel in Figure 4.3).

For all probe types (i.e., CE, DMC, and D-Cue), the participant and researcher looked at the laptop screen, the researcher 'read' the relational network to/with the participant, and then 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. All probe sets included five trials, and passing criteria required responding correctly on all five trials (100%) the first time the probe was presented or scoring 100% correct twice consecutively. CE Probe Set 1 was presented in baseline sessions, and it was the first probe set after training commenced, followed by novel CE Probes, DMC Probes, novel CE Probes again, and D-Cue Probes, in that order. Failed D-Cue Probes were not followed by more training sessions. Maintenance probes for all three probe types were administered one month after each participant's last probe session.

The training condition itself included two phases (Training Phase 1 and Phase 2). A modified version of the Greer and Ross (2008) decision-making protocol was followed during the training condition. If Phase 1 training data showed three increasing data points, then training in Phase 1 continued and was re-evaluated at five data points. If Phase 1 data showed three stable or three descending data paths, then Phase 2 training would be implemented. One participant required Phase 2 training.

Training sets included the same relational network structure as in the CE Probes (i.e., [A [S] B], [B [S] C], [C [D] D]), but with novel stimuli (i.e., different coloured monochromatic circles). Each training set included two PowerPoint slides; the first slide showed the relational network at the top of the screen, and the second slide showed the analogical stimuli below the relational network. Upon presentation of the relational network (the first slide), the researcher and participant read the relational network together. After reading the network, the second slide with the analogical stimuli was presented. First, the researcher asked the participant to derive the combinatorially entailed (CE) relation for the stimuli in each compound element (the sample and the two comparisons) by pointing to one of the compound elements and asking, 'Is A the same or different to D?' Correct responses included identifying the relation between the stimuli in the compound by saying 'same' or 'different'. Feedback was provided for each response; for correct responses, verbal praise and stickers were immediately delivered, and for incorrect responses, the researcher provided the correct response and re-presented the question.

Following the three CE relational questions, the researcher presented the instruction for the analogy task, 'Now look at this one at the top (points to the sample), which one of these (points to comparisons) is like the one at the top?' Correct responses included either pointing to the correct comparison compound element or vocally stating the two colours of the correct comparison compound element. Correct responses were consequated with verbal praise and a sticker, and CE Probe Set 1 was re-presented. Incorrect training responses were followed by a correction procedure, including: 1) repeating the combinatorially entailed relational questions for all compound elements in the analogy, and 2) re-presenting the analogical question plus pointing to the correct comparison compound element and saying, 'this one is like the one at the top.' Following either a correct or incorrect trial, the CE Probe was re-presented. If the participant met the criteria for a phase change, Phase 2 training was implemented. Phase 2 training followed the same training procedure as in Phase 1 but included three training trials.

Interobserver Agreement

Procedural fidelity checks and interobserver agreement (IOA) were determined for baseline, probes, and training conditions by a trained research assistant. Procedural fidelity was assessed using 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 16% of all trials and was 100%. Interobserver agreement was calculated on a trial-by-trial basis for each probe and training trial. Interobserver agreement was assessed for 23% of all trials and was 100%.

Results

Pre-Assessment

Ten children were given the pre-assessment measure. Six of these children did not pass the relational section of this test and thus did not proceed to the analogical section of the test. Of the four children who passed the relational section, three of those went on to take part in the multiple baseline training.

Table 4.2

Pre-Assessment Relational and Analogical Scores

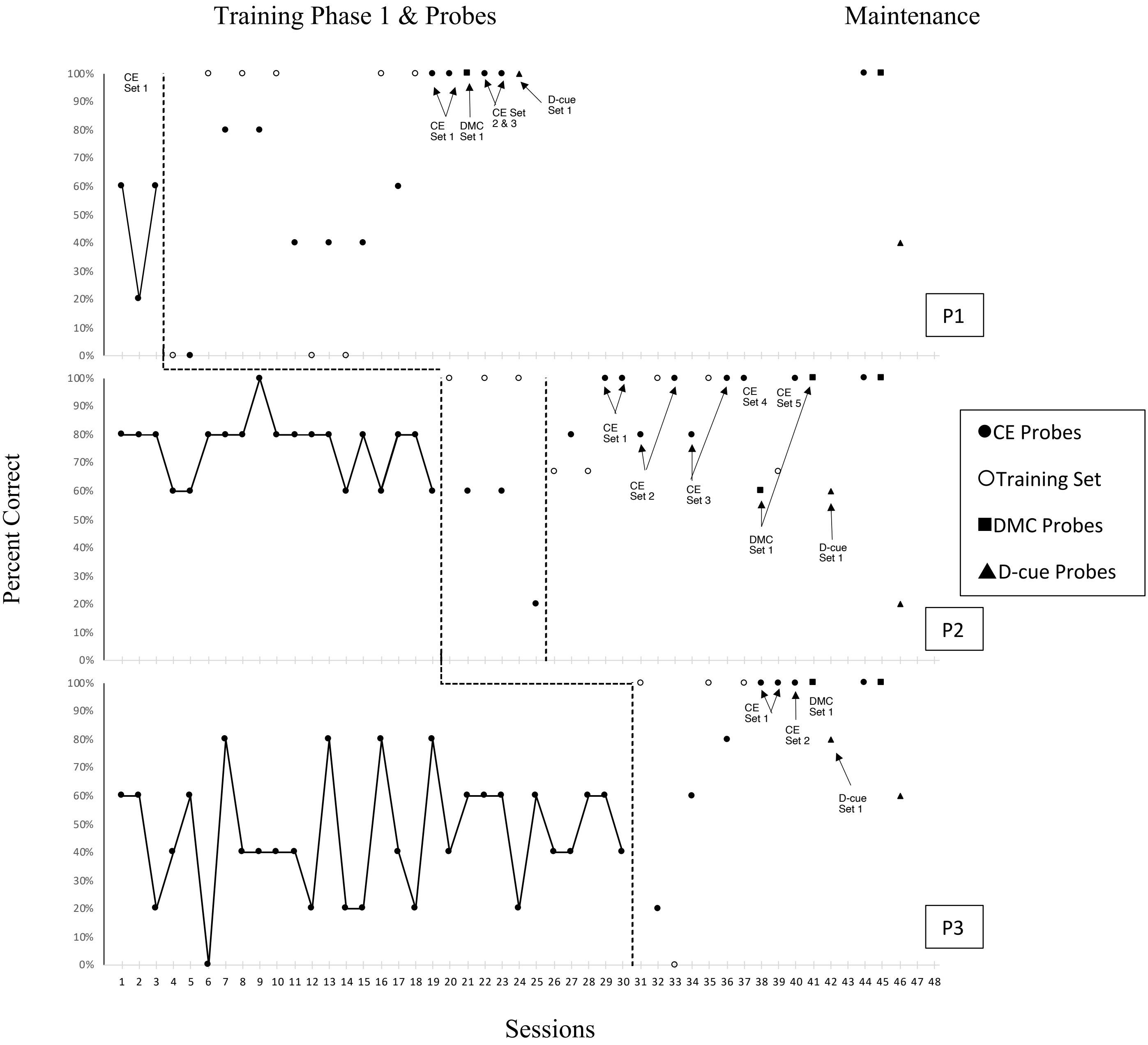
Participant	Relational Score	Prompt Level 1		Prompt Level 2	Prompt Level 3	Inclusion Criteria: Mean Scores	
		Relational	Analogy	Analogy	Analogy	Relational	Analogy
	%	%	%	%	%	*M*	*M*
P1	86	93	20	60	40	89.5	40
P2	100	87	20	40	40	93.5	33
P3	93	87	60	60	40	90	53

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

In what follows, we describe in more detail the pre-assessment performance of the three children who went on to be trained in the multiple baseline training (see Table 4.2 for a summary of participant pre-assessment performance). P1 scored 12/15 (86%) correct on the last set of combinatorial entailment and compound identification trials, 14/15 (93%) correct on the combinatorial entailment relational section of prompt Level 1, and 1/5 (20%) correct on the analogical section of prompt Level 1; 3/5 (60%) on prompt Level 2; and 2/5 (40%) on prompt Level 3. Hence P1 obtained 89.5% correct on the relational section and 40% correct on the analogical section of the pre-assessment measure. P2 scored 14/14 (100%) correct on the last set of combinatorial entailment and compound identification trials, 13/15 (87%) correct on the combinatorial entailment relational section of prompt Level 1, and 1/5 (20%) on the analogical section of prompt Level 1; 2/5 (40%) on prompt Level 2; and 2/5 (40%) on prompt Level 3. Hence P2 obtained 93.5% correct on the relational section and 33% correct on the analogical section of the pre-assessment measure. P3 scored 13/14 (93%) correct on the last set of combinatorial entailment and compound identification trials, 13/15 (87%) correct on the combinatorial entailment relational section of prompt Level 1, and 3/5 (60%) on the analogical section of prompt Level 1; 3/5 (60%) on prompt Level 2; and 2/5 (20%) on prompt Level 3. Hence P3 obtained 90.5% correct on the relational section and 33% correct on the analogical section of the pre-assessment measure.

Figure 4.4

Participant Responding in Baseline and Training Conditions



Note. CE Probes: Combinatorially entailed analogical responses; DMC Probes: Trained, mutually entailed, and combinatorially entailed analogical responses; D-Cue Probes: The contextual cue for difference varied in location within the relational network.

Training

In the multiple baseline training condition, P1 scored 0% correct on his first training session and scored 0% on the following CE Probe Set 1 (see Figure 4.4). After eight training sessions, P1 scored 100% on CE Probe Set 1 twice consecutively, 100% correct on the first generalisation probe, DMC Probe Set 1, 100% correct on two novel CE Probes (CE Set 2 and Set 3), and 100% correct on the D-Cue Probe. In maintenance testing, P1 scored 100% on the CE Probe (Set 4), 100% on the DMC Probe, and 40% on the D-Cue Probe.

P2 scored 100% correct on her first training session and 60% on the subsequent CE Probe. After three training sessions, P2 scored 20% on CE Probe Set 1 and Phase 2 training was implemented. P2 scored 67% correct on the first Phase 2 training and set and her CE Probe Set 1 score increased to 80% correct. P2 required one more training set before scoring 100% twice consecutively on CE Probe Set 1. P2 scored 80% correct on CE Probe Set 2 and required one training set to score 100% correct on CE Probe Set 2. P2 scored 80% correct on CE Probe Set 3 and required one training set to score 100% correct on CE Probe Set 3. P2 scored 100% correct on CE Probe sets 4 and 5, 100% correct DMC Probe Set 1, and 60% correct on the D-Cue Probe. At one-month maintenance testing, P2 scored 100% on the CE Probe (Set 6), 100% on the novel DMC Probe, and 20% on the D-Cue Probe.

P3 scored 100% correct on his first training session and scored 60% on the following CE Probe Set 1. After four training sessions, Participant P3 scored 100% on CE Probe Set 1 twice consecutively and 100% on CE Probe Set 2. Participant P3 scored 80% correct on DMC Probe Set 1 and required three training sessions before scoring 100% twice consecutively on DMC Probe Set 1 and 100% on a novel DMC Probe. Following the DMC Probes, P3 scored 100% on the maintenance CE Probe Set 3, and 60% on the D-Cue Probe. During the one-month follow-up maintenance testing, P3 scored 100% correct on all three probes.

P3 scored 100% correct on his first training session and scored 20% correct on the following CE Probe Set 1. After three more training sessions, Participant P3 scored 100% on CE Probe Set 1 twice consecutively and 100% on CE Probe Set 2. Participant P3 scored 100% correct on DMC Probe Set 1 and 80% correct on the D-Cue Probe. During the one-month follow-up maintenance testing, P3 scored 100% on the CE Probe and DMC Probe, and 60% correct on the D-Cue Probe.

Summary

Three out of ten children given pre-assessment testing met criteria and hence were admitted into multiple baseline training. Following analogy training, all three successfully showed analogical responding during CE Probe sets, including the original CE Probe Set 1 used during baseline testing, as well as a novel CE Probe Set 2. All participants scored 100% correct on the DMC Probes. Only P1 scored 100% correct on the D-Cue Probe. At one-month maintenance probes, all three participants scored 100% correct on a novel CE Probe Set and a novel DMC Probe Set, and all three participants failed the D-Cue Probe.

Discussion

After direct training in relating combinatorially entailed relations, all three participants in Study 2 showed analogical responding according to RFT's conception of analogy as the derived relating of relations. This extends previous work by Carpentier et al. (2002, 2003), who were the first to use this RFT approach to examine analogy in young children. In the Carpentier et al. studies, 5-year-old children initially failed to derive equivalence-equivalence relations without additional prompting. More specifically, the 5-year-old children required pre-testing with compound-compound-matching tasks involving trained relations (i.e., not derived relations) before they successfully passed the derived compound relations (BC-BC) test. These results indicated that 5-year-old children are capable of analogical reasoning, but in this case, they required prompting before doing so. The present data support the Carpentier et al. findings that 5-year-old children are capable of analogical responding. However, the present study extends Carpentier et al. Rather than simply facilitating the emergence of analogy through prompting procedures, the present data extend this previous research by directly training analogy through multiple exemplar training using a controlled multiple baseline design such that the participants of the study could thereafter demonstrate analogical responding without additional prompting procedures being needed.

The results from the multiple baseline showed that the CE analogy training procedure was an effective intervention for training analogy and eliciting generative responding, as shown by the generalisation data. All participants passed baseline CE Probe Set 1 as well as novel CE Probe sets following CE analogy training. Furthermore, correct analogical responding generalised to the DMC Probes. Previous RFT work on relating relations (e.g., Barnes et al., 1997; Carpentier et al., 2002) typically involved testing and training at least two sets of combinatorially entailed relations. The present work included directly presented, mutually entailed, and combinatorially entailed relations to provide evidence that participants were generalising analogical responding across different types of relational derivation rather than limiting their responses to the CE relation only.

Only Participant P1 scored 100% correct on the first presentation of the D-Cue generalisation probes. Participants P2 and P3 scored 60% and 80%, respectively. During one-month follow-up maintenance testing, all participants scored 100% correct on the novel CE Probes and the novel DMC Probes, and none of the participants passed the novel D-Cue Probe. In the D-Cue Probes, the format of the relational network did not follow the [A [S] B], [B [S] C], [C [D] D] pattern. There were always two sameness cues and one difference cue in the relational networks, but across D-Cue Probe trials, the difference cue could be the first, second, or third relational cue. As in the CE Probes, combinatorial entailment was required to derive the relation between the two stimuli in all sample and comparison compound elements. It is unclear why participants did not score higher on the D-Cue Probes. A limitation of the present study is that additional training was not provided following the failed probe. The D-Cue Probe was used only to examine whether generalisation to such a probe might be seen; it remains unknown if participants might have passed the D-Cue Probes with more CE analogy training. Future researchers could investigate whether generalisation would eventually extend to the D-Cue Probe (as well as other novel probes) with more CE analogy training. Study 2 of the present thesis indicated that 5-year-old participants might not have had sufficient practice with arbitrary relational responding to relate relations successfully. Thus, more practice with multiple exemplars in relational responding may be necessary for generalisation to occur in more complex relational responding such as in relating relations. Furthermore, the present training procedure may be found to be more effective if participants were required to respond correctly before the CE Probe was re-presented. In the present study, the CE Probe was re-presented whether the participant emitted a correct or incorrect response during the correction procedure. The next study in this thesis, Study 3, is a replication of the present study with a modified correction procedure.

Chapter 5: Study 3. Testing and Training Analogy in Young Children Part 2: Improving the Training Protocol

Portions of this chapter have been published:

Kirsten, E. B., Stewart, I., & McElwee, J. (2021). Testing and Training Analogical Responding in Young Children Using Relational Frame Theory. *The Psychological Record*. https://doi.org/10.1007/s40732-021-00468-9

Introduction

The present study sought to extend Study 2 (Chapter 4) and further examine the acquisition of analogical responding operationalized as derived relations between relations in five-year old children. In Study 2, three five-year old children were assessed and trained in relating relations using an RFT-based protocol in a multiple baseline design. Following multiple exemplar training, correct responding increased to criterion levels for all three children, and both generalization and maintenance were observed; all three participants showed analogical responding according to RFT's conception of analogy as the derived relating of relations after direct training in relating combinatorially entailed relations in a multiple baseline design. Study 2 extended previous work by Carpentier et al. (2002, 2003), who were the first to use this RFT approach to examine analogy in young children.

The results from the multiple baseline showed that the CE analogy training procedure was an effective intervention for training analogy and eliciting generative responding, as shown by the generalisation data. All participants passed baseline CE Probe Set 1 as well as novel CE Probe sets following CE analogy training.

A possible limitation noted in Study 2 was that the training procedure may be rendered more effective if participants were required to respond correctly before the CE Probe was presented. In Study 2, the CE Probe was presented whether the participant emitted a correct or incorrect response during the training correction procedure. In the present study, Study 3, a modified correction procedure was implemented. Participants were required to respond correctly in the training trials before the CE Probe was re-presented. That is, in the present study incorrect training responses were followed by a correction procedure including re-presentations of the combinatorially entailed relational questions for all compound elements in the analogy and the researcher modeled the correct response for the analogy trial. Following an incorrect trial, training trials were presented until the participant emitted a correct response, and then the CE Probe was re-presented.

Method

Participants and Setting

Participants were three typically developing children enrolled in a private elementary school on the east coast of the United States. P1 was a female aged 5 years and 11 months, P2 was a male aged 5 years and 7 months, and P3 was a male aged 5 years and 9 months. Participants were selected for inclusion based on their performance on an adaptation of the relational assessment used in Study 1 of the present thesis (see Figure 5.1 for a summary version of the pre-assessment). The researcher administered all probe and training sessions in an otherwise unoccupied classroom of the participants' school during school hours. The researcher sat next to the participant at a standard school desk. A second independent observer sat approximately three 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 the 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 consent was obtained from each participant.

Experimental Design

As in Study 2, a combination of multiple baseline design across participants and a multiple probe design across behaviours was used in this study. The study included a relational pre-assessment for screening potential participants; a baseline condition in which analogy was tested; and a training condition in which analogical responding was trained, and generalisation probe trials were presented. For P2 to enter the training condition, the participant previously exposed to training (i.e., P1) had to meet the training criterion (i.e., scoring 100% correct twice consecutively on the trained repertoire) while for P3 to enter the training condition, P2 had to meet the training criterion.

Materials & Apparatus

The same materials and technological equipment were used in the present study as in Study 2.

Procedure

Overview

The pre-assessment was administered first in one session and took between 5 and 25 minutes to complete. Following the pre-assessment, participants entered the multiple baseline, including baseline sessions followed by training and multiple probe sessions. Sessions lasted between 10 and 40 minutes, and on average, included seven probe and training sets; a probe set took approximately two to four minutes to complete, and a training set took approximately one to three minutes to complete. All participants started baseline sessions at the same time, and all participants took approximately three weeks to complete the study. Maintenance probe sessions were administered one month after each participant's last probe in the training condition was administered. See Table 5.1 for a schematic presentation of the pre-assessment and multiple baseline conditions used in Study 2 and Study 3 of this thesis.

Multiple Baseline Conditions

As in Study 2, the multiple baseline components included a baseline condition including unreinforced probe sessions and a training/testing condition in which training and probe sessions were administered. The study included the same three types of probes, including Combinatorially Entailed Analogy Probes (CE Probes); Directly Presented/Mutually Entailed/Combinatorially Entailed Analogy Probes (DMC Probes); and Difference-Cue-Shifted Analogy Probes (D-Cue Probes) (see Figure 5.3 for an illustrative example of each probe type).

During the training condition, the three CE relational questions preceded the instruction for the analogy task, 'Now look at this one at the top (points to the sample), which one of these (points to comparisons) is like the one at the top?' Correct responses included either pointing to the correct comparison compound element or vocally stating the two colours of the correct comparison compound element. Correct responses were consequated with verbal praise and a sticker, and CE Probe Set 1 was re-presented. In Study 3, incorrect training responses were followed by a correction procedure, including: 1) repeating the combinatorially entailed relational questions for all compound elements in the analogy, and 2) re-presenting the analogical question plus pointing to the correct comparison compound element and saying, 'this one is like the one at the top.' Following an incorrect trial, another training trial was presented until the participant emitted a correct response, and then the CE Probe was re-presented. (In the previous study, Study 2, the CE Probe was re-presented after only one trial, regardless of the participant's response.) If the participant failed the CE Probe, another training set was presented, and the training and probing procedure was repeated until all probe types had been administered.

Table 5.1

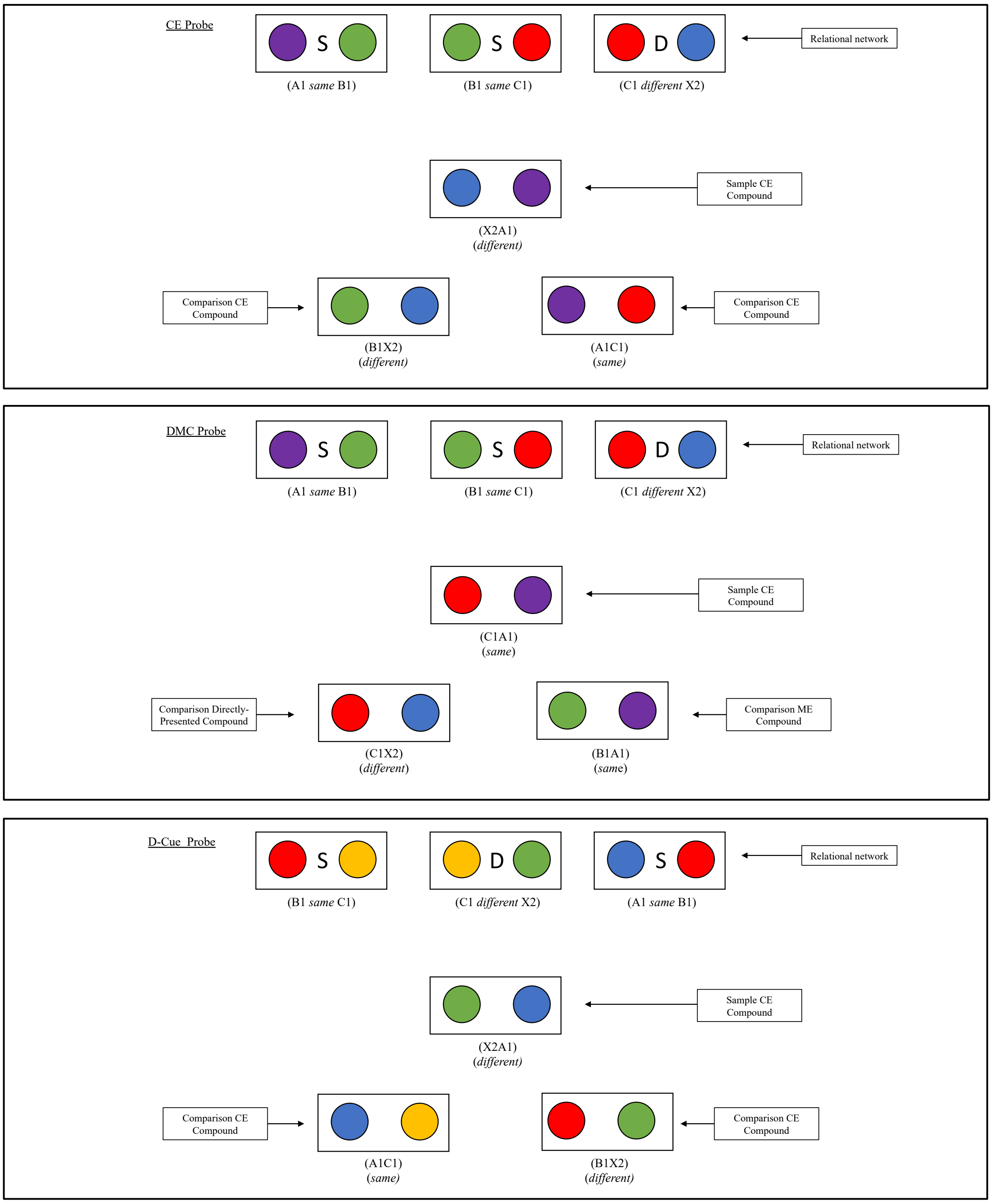
Schematic Presentation of the Pre-Assessment and Multiple Baseline Condition

	Phase	Relational Targets	Passing Criteria %
Pre-Assessment			
	1.	Identifying contextual cues	100
	2.	Reading relational networks	100
	3.	Testing DP and ME relations	100
Inclusion Criteria	4.	Testing CE relations	80
	5.	Testing compound elements	80
		Analogy Probes	
	6.	Level 1: 3 CE trials + 5 analogy trials	100
	7.	Level 2: 5 analogy trials	100
	8.	Level 3: 5 analogy trials	100
Multiple Baseline Study			
	Baseline: CE Probe Set 1		100
Training Condition & Probes	•	Training: 3 CE trials + 1 analogy trial	100
	•	CE Probe Set 1: 5 trials	100
	•	Novel CE Probe: 5 trials	100
	•	DMC Probe: 5 trials	100
	•	Novel CE Probe: 5 trials	100
	•	D-Cue Probe: 5 trials	100
One-Month Maintenance Probes		Novel CE Probe	100
		Novel DMC Probe	100
		Novel D-Cue Probe	100

Note. DP = directly presented relations; ME = mutually entailed relations; CE = combinatorially entailed relations.

Figure 5.1

Three Analogy Probe Types: CE Probes, DMC Probes, and D-Cue Probes



Note. First panel: combinatorially entailed relations; second panel: mixed directly presented, mutually entailed, and combinatorially entailed relations; third panel: the difference cue changes location in the relational network plus combinatorially entailed relations.

Interobserver Agreement

Procedural fidelity checks and interobserver agreement (IOA) were determined for baseline, probes, and training conditions by a trained research assistant. Procedural fidelity was assessed using 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%. Interobserver agreement was calculated on a trial-by-trial basis for each probe and training trial. Interobserver agreement was assessed for 43% of all trials and was 100%.

Results

Pre-Assessment

Eight children were given the pre-assessment measure. Four of these children did not pass the relational section of this test and thus did not proceed to the analogical section of the test. Of the four children who passed the relational section, three of those went on to take part in the multiple baseline training.

Table 5.2

Pre-Assessment Combinatorially Entailed Relations and Analogy Scores

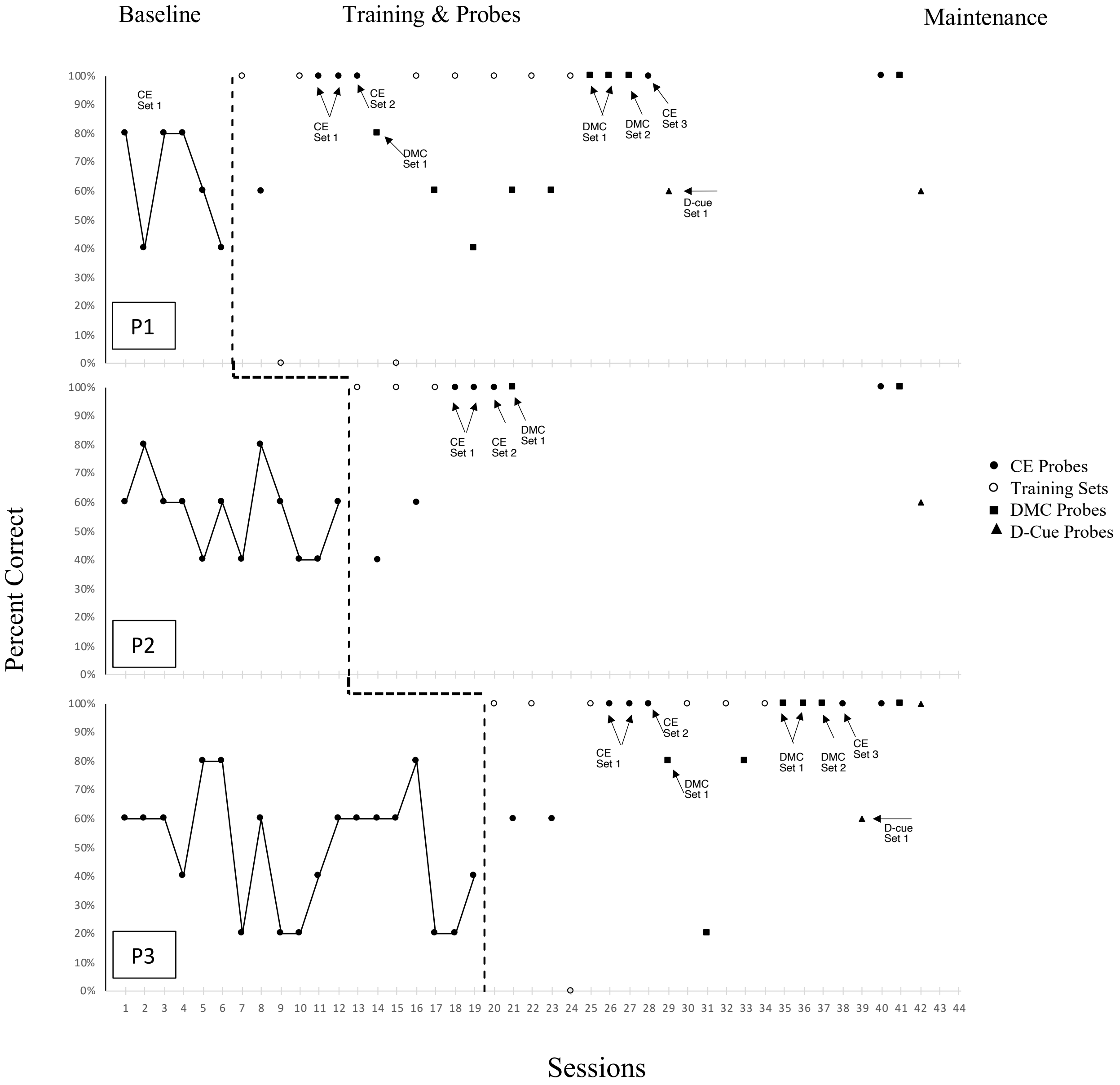
Participant	Relational Score	Prompt Level 1		Prompt Level 2	Prompt Level 3	Inclusion Criteria: Mean Scores	
	%	Relational %	Analogy %	Analogy %	Analogy %	Relational *M*	Analogy *M*
P1	79	93	60	40	40	86	47
P2	100	100	40	0	20	100	20
P3	100	93	40	40	20	96.5	33

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

In what follows, we describe in more detail the pre-assessment performance of the three children who went on to be trained in the multiple baseline training (see Table 5.2 for a summary of participant pre-assessment performance). P1 scored 11/14 (79%) correct on the last set of combinatorial entailment and compound identification trials, 14/15 (93%) correct on the combinatorial entailment relational section of prompt Level 1, and 3/5 (60%) correct on the analogical section of prompt Level 1; 2/5 (40%) on prompt Level 2; and 2/5 (40%) on prompt Level 3. Hence P1 obtained 86% correct on the relational section and 47% correct on the analogical section of the pre-assessment measure. P2 scored 14/14 (100%) correct on the last set of combinatorial entailment and compound identification trials, 15/15 (100%) correct on the combinatorial entailment relational section of prompt Level 1, and 2/5 (40%) on the analogical section of prompt Level 1; 0/5 (0%) on prompt Level 2; and 1/5 (20%) on prompt Level 3. Hence P2 obtained 100% correct on the relational section and 20% correct on the analogical section of the pre-assessment measure. P3 scored 14/14 (100%) correct on the last set of combinatorial entailment and compound identification trials, 14/15 (93%) correct on the combinatorial entailment relational section of prompt Level 1, and 2/5 (40%) on the analogical section of prompt Level 1; 2/5 (40%) on prompt Level 2; and 1/5 (20%) on prompt Level 3. Hence P3 obtained 96.5% correct on the relational section and 33% correct on the analogical section of the pre-assessment measure.

Figure 5.2

Participant Responding in Baseline and Training Sessions



Note. CE Probes: Combinatorially entailed analogical responses; DMC Probes: Trained, mutually entailed, and combinatorially entailed analogical responses; D-Cue Probes: The contextual cue for difference varied in location within the relational network.

Training

In the multiple baseline training condition, P1 scored 100% correct on her first training session and scored 60% on the following CE Probe Set 1. After three training sessions, P1 scored 100% on CE Probe Set 1 twice consecutively and 100% on CE Probe Set

2. P1 scored 80% correct on the first generalisation probe session, DMC Probe Set 1. P1's responses decreased to 0% correct during the training session after the first DMC Probe, and she required five more training sessions before scoring 100% on two consecutive DMC Set 1 Probes and one novel DMC Probe. On the first maintenance check, P1 scored 100% on CE Probe Set 3. P1 scored 60% on the D-Cue Probe. In maintenance testing, P1 scored 100% on the CE Probe, 100% on the DMC Probe, and 60% on the D-Cue Probe (see Figure 5.4).

P2 scored 100% correct on his first training session and 40% on the subsequent CE Probe. After three training sessions, P2 scored 100% on CE Probe Set 1 twice consecutively and 100% on CE Probe Set 2. P2 scored 100% on the first DMC Probe and hence did not require another DMC Probe. P2 subsequently began to exhibit behavioural issues that interfered with his participation, and as a result, he did not complete the D-Cue Probe after training. However, he completed all three probe types at one-month maintenance testing and scored 100% on CE Probe Set 3, 100% on the novel DMC Probe, and 60% on the D-Cue Probe.

P3 scored 100% correct on his first training session and scored 60% on the following CE Probe Set 1. After four training sessions, Participant P3 scored 100% on CE Probe Set 1 twice consecutively and 100% on CE Probe Set 2. Participant P3 scored 80% correct on DMC Probe Set 1 and required three training sessions before scoring 100% twice consecutively on DMC Probe Set 1 and 100% on a novel DMC Probe. Following the DMC Probes, P3 scored 100% on the maintenance CE Probe Set 3 and 60% on the D-Cue Probe. During the one-month follow-up maintenance testing, P3 scored 100% correct on all three probes.

Summary

Three out of eight children given pre-assessment testing met criteria and hence were admitted into multiple baseline training. Following analogy training, all three successfully showed analogical responding during CE Probe sets, including the original CE Probe Set 1 used during baseline testing, as well as a novel CE Probe Set 2. All participants scored 100% correct on the DMC Probes. P1 and P3 both scored 60% correct on the D-Cue Probes, while P2 did not complete the D-Cue Probe following training. At one-month maintenance probes, all three participants scored 100% correct on a novel CE Probe Set and a novel DMC Probe Set, while all participants, including P2, scored only 60% correct on the D-Cue Probe.

Discussion

After direct training in relating combinatorially entailed relations, all three participants in Study 3 showed analogical responding according to RFT’s conception of analogy as the derived relating of relations. As in Study 2, this extends previous work by Carpentier et al. (2002; 2003) in which 5-year-old children initially failed to derive equivalence-equivalence relations without additional prompting. The data from Studies 2 and 3 support the Carpentier et al. findings that 5-year-old children are capable of analogical responding. However, the present studies extend Carpentier et al. Rather than simply facilitating the emergence of analogy through prompting procedures, the present data extend this previous research by directly training analogy through multiple exemplar training using a controlled multiple baseline design such that the participants of the study could thereafter demonstrate analogical responding without additional prompting procedures being needed.

The results from the multiple baseline showed that the CE analogy training procedure was an effective intervention for training analogy and eliciting generative responding, as shown by the generalisation data. All participants passed baseline CE Probe Set 1 as well as novel CE Probe sets following CE analogy training. Furthermore, correct analogical responding generalised to the DMC Probes. Previous RFT work on relating relations (e.g., Barnes et al., 1997; Carpentier et al., 2002) typically involved testing and training at least two sets of combinatorially entailed relations. The present work included directly presented, mutually entailed, and combinatorially entailed relations to provide evidence that participants were generalising analogical responding across different types of relational derivation rather than limiting their responses to the CE relation only.

As in Study 2, none of the participants scored 100% correct on the first presentation of the D-Cue generalisation probes. However, during one-month follow-up maintenance testing, one participant scored 100% correct on all probe types, including the D-Cue Probe. As discussed in Study 2, in the D-Cue Probes, the format of the relational network did not follow the [A [S] B], [B [S] C], [C [D] D] pattern. There were always two sameness cues and one difference cue in the relational networks, but across D-Cue Probe trials, the difference cue could be the first, second, or third relational cue. As in the CE Probes, combinatorial entailment was required to derive the relation between the two stimuli in all sample and comparison compound elements. It is unclear why participants did not score higher on the D-Cue Probes. A limitation of the present study is that additional training was not provided following the failed probe. The D-Cue Probe was used only to examine whether generalisation to such a probe might be seen; it remains unknown if participants might have passed the D-Cue Probes with more CE analogy training. Future researchers could investigate whether generalisation would eventually extend to the D-Cue Probe (as well as other novel probes) with more CE analogy training. In Study 1, the data indicated that 5-year-old participants might not have had sufficient practice with arbitrary relational responding to relate relations successfully. Thus, more practice with multiple exemplars in relational responding may be necessary for generalisation to occur in more complex relational responding such as in relating relations. Using the REP format, future researchers could examine the effects of additional arbitrary relational training on analogical responding.

Previous RFT studies have successfully shown that many different patterns of relational framing can be tested and trained via multiple exemplar training (e.g., Belisle et al., 2016; Berens & Hayes, 2007; Cassidy et al., 2011, 2016; Mulhern et al., 2018). These include not just simple frames such as comparison and opposition (e.g., Berens & Hayes, 2007; Cassidy et al., 2011, 2016) but also relatively more complex patterns of framing such as hierarchy (Mulhern et al., 2018) and deixis (Belisle et al., 2016). The results from the present study have provided evidence that derived relating of relations can also be trained using multiple exemplar training just like the others listed, and as such, this is yet further evidence in support of the RFT thesis. This work might also hopefully prompt further work on testing and training analogical relations in young children, including children diagnosed with ASD and other language delays. For example, considering that analogy seems centrally important for language and cognition, training analogy in children struggling with language development could result in significant language generativity in addition to facilitating intellectual growth. The closest behavioural study of analogy-type relations in children since Carpentier et al. is on metaphorical responding in children diagnosed with ASD (Persicke et al., 2012). Considering the prevalence of figurative language in our socio-verbal environment, Persicke et al. argued, children with ASD face significant comprehension challenges due to their difficulty understanding non-literal language, including metaphor. Using multiple exemplars of metaphors, Persicke et al. successfully taught metaphorical language to all three participants. In addition, the results showed generalisation to untrained, novel metaphors. These data are promising for children with ASD who have mastered the basic verbal operants (i.e., echoics, mands, tacts, and intraverbals) and are ready for more complex language training. However, Persicke et al. noted methodological issues worth considering in future research; the relative difficulty of the exemplar metaphors was not quantified; thus, the level of difficulty was not equal across metaphors. Secondly, the researchers could not control for participant history with the metaphors, which could have affected the training and testing results.

The present data also constitute an important addition to the literature on analogical relations beyond behaviour analysis. Cognitive researchers have studied the acquisition of analogical responding in young children since Piaget's early work in the 1970s (e.g., Piaget et al., 1977/2001; Goswami, 1989, 1996, 2001; Rattermann & Gentner, 1998; Richland et al., 2006). However, as has been argued in previous RFT literature, this approach to analogical development has failed to lead to a functional analytic model of this behaviour that might lend itself to training this repertoire. Moreover, within the cognitive literature, the age at which analogical responding develops in young children remains a debate amongst researchers (Richland et al., 2006), and flaws in testing methodologies have hindered progress in testing and training analogy in this population. For example, Goswami and Brown (1989) provided data of analogical responding in 3-year-old children. However, Rattermann and Gentner (1998) provided similarity ratings for the stimuli used to test for analogy in Goswami and Brown (1989) and concluded that children younger than five responded correctly based on object similarity rather than analogical relations. Another vein of cognitive research has examined the use of analogy to facilitate learning in other domains, but training analogical responding itself was not included (e.g., Brown & Kane, 1988; Polya, 1945/2004; Richland et al., 2010). Thus, within the cognitive literature, the development of analogical responding in young children has been theorised and examined, and the use of analogies to solve daily problems encountered in life or at school has been examined, but there is a dearth of experimental work on the effective training of analogy in young children. In contrast, RFT's functional analytic definition of analogy has allowed a reliable methodology for targeting analogical reasoning in this population.

One important feature of Studies 2 and 3 is that analogical responding was tested and trained using a novel RFT-based protocol, and more specifically, an adaptation of the RFT-based protocol used in Study 1 of the present thesis. The utilisation of shapes across all trials obviated the need to control for participants' previous experience and knowledge as well as controlled for stimulus consistency across trials. The results of this study suggest the applied and experimental potential of this format. For example, by using monochromatic circles and single-lettered contextual cues, the REP format makes it easy to produce multiple training and testing sets suitable for young children and pre-readers, which, in turn, makes it possible to directly train children on complex relations with many exemplars. Moreover, this format does not require any relational pre-training but instead allows for faster testing (with multiple versions of tests) than is possible with the tricky methodological issues in MTS procedures. Finally, a participant's previous experience and knowledge is not an experimental confound. These favourable variables of the REP format permitted a multiply controlled study in which we could target analogy testing and training directly while maintaining experimental control. That is, the REP format afforded us with quick and effective stimulus control allowing us to implement a multiple baseline design to examine the efficacy of multiple exemplar training to establish the core repertoire. For example, once participants were trained on the format of the REP in the pre-assessment, we were able to immediately implement analogical baseline testing across all participants and then implement controlled training and testing conditions with novel sets of stimuli. In addition, the REP format provided design flexibility for different generalisation probe types. In contrast, the time-consuming and laborious pre-training required in match-to-sample procedures, for example, would have made multiple testing and training conditions in a controlled multiple baseline design unfeasible.

Future researchers could increase the number of stimuli in the relational network to include more than four relata (i.e., the stimuli being related). A larger array would more closely replicate the analogical work by Carpentier et al. (2002) in which participants were trained to identify nine arbitrary pictures as being equivalent or non-equivalent and were then tested for equivalence-equivalence and non-equivalence-non-equivalence (i.e., relating relations). Future researchers could introduce multiple relational networks by increasing the array of monochromatic circles and using only the sameness cue, S, to relate particular circles to one another. For example, as in Carpentier et al., an array of nine circles could be divided into three, three-member sets; participants would refer to the relational networks to determine equivalence or non-equivalence between compound elements, followed by equivalence-equivalence or non-equivalence-non-equivalence testing and training. A larger array would allow for more combinatorially entailed sample and comparison compound combinations, thus circumventing a potential limitation in the present study. That is, the relational networks were composed of two sameness cues and one difference cue (e.g., Red S Blue, Blue S Green, Green D Yellow) and therefore permitted only combinatorially derived difference relations. Thus, the compound elements in the CE and D-Cue Probe trials had to be difference relations. Sameness relations were possible via directly presented, mutually entailed, or combinatorially entailed relations as in the DMC Probes. It is possible that participants learnt to pick the difference comparison compound instead of relating difference-difference relations. However, during the DMC Probes, it was possible for both sameness and difference relations to be tested because the stimuli in the sample and comparison compound elements included directly trained, mutually entailed, and combinatorially entailed relations. All participants passed the DMC Probes, which suggests that they were responding to analogical relations rather than just selecting the difference compound. However, future work could further examine testing and training analogical relations as relating relations by increasing the number of relational networks.

The REP format could also be extended to test and train analogical relations beyond coordination and distinction. Previously, Lipkens and Hayes (2009) successfully looked at multiple relations (i.e., sameness, difference, comparison, and opposition) in analogy in adult participants. Lipkens and Hayes tested across selection-based and topography-based tasks, including selecting the correct relata, selecting the correct relational cue, and producing the correct relata. However, the Lipkens and Hayes procedure required extensive pre-training as well as reading and writing skills. For young children, it might be possible to test and train multiple relations and topography-based responses using variations of the REP format used in the present study. For example, in Lipkens and Hayes' selecting relata task, the participant was given the sample, a relational cue, and two comparisons; participants had to select the correct comparison based on the relational cue; in the selecting the correct relation task, the participant had to select the correct relational cue given the sample and the comparison stimulus; and in the producing relata task, the participant had to produce the correct relata given the sample stimulus and relational cue. The REP format would support these various tasks and make them accessible to young children. For example, in producing the relata tasks, the response could include shading in a blank circle with the correct colour, and in selecting the relation tasks, the same label format (i.e., S or D) used in the current study could be applied and extended to other relations (see for example Study 1 in Chapter 3).

Future researchers could also examine the effects of training sameness relations on the emergence of other relations. For example, once participants have been trained in coordinate analogical responding, future researchers could probe other types of relations (e.g., comparative-comparative) for generalisation. The results in Study 1 showed that coordinate analogical responding was acquired before comparative, opposite, temporal, and hierarchical analogical responding. Thus, future research could also use the present format to test and train children to relate these other relations.

Considering the relevance of analogy to intelligence, future researchers could investigate the generalised effects of training analogical responding on socially valid measures. For example, future dependent variables could include mainstream analogy tests, academic achievement tests, or standardised tests of cognitive performance. Furthermore, since relating relations allows response classes to affect other response classes, future researchers could also investigate whether training analogical responding in young children facilitates the emergence of developing repertoires such as reading or mathematics.

Studies 2 and 3 add to the limited behavioural research on analogical responding in young children and contribute further evidence that 5-year-old children can be trained in analogical responding. Moreover, by offering a precise, functional analytic model of analogy, the present study has arguably shed additional light on this phenomenon beyond that provided by mainstream cognitive psychological work. This work provides further evidence that analogy tends to be acquired around the age of five but goes beyond previous work in showing how multiple exemplar training of the core pattern involved might accelerate this acquisition. Considering the ubiquity of analogical responding in everyday life, more research regarding its development and training in young children both with and without 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 research and applied utility.

Chapter 6: Study 4. Testing and Training Analogy in Young Children Part 3: (i) A More Comprehensive Model (ii) Testing Children with ASD

Portions of this chapter have been accepted for publication:

Kirsten, E. B., Stewart, I., & McElwee, J. (2022). Testing and Training Analogical Relational Responding in Children with and Without Autism. *The Psychological Record*. https://doi.org/10.1007/s40732-021-00493-8

Introduction

In the present study, we modified the testing and training procedure used in Studies 2 and 3 to include a larger array of stimuli as suggested in the previous study, Study 3. A larger array would allow for more combinatorially entailed sample and comparison compound combinations, thus circumventing a potential limitation in Study 3. That is, the relational networks were composed of two sameness cues and one difference cue (e.g., Red S Blue, Blue S Green, Green D Yellow) and therefore permitted only combinatorially derived difference relations. Thus, the compound elements in the CE and D-Cue Probe trials had to be difference relations. Sameness relations were possible via directly presented, mutually entailed, or combinatorially entailed relations as in the DMC Probes. It is possible that participants learnt to pick the difference comparison compound instead of relating difference-difference relations.

A larger array would more closely replicate the analogical work by Carpentier et al. (2002) in which participants were trained to identify nine arbitrary pictures as being equivalent or non-equivalent and were then tested for equivalence-equivalence and non-equivalence-non-equivalence (i.e., relating relations). Future researchers could introduce multiple relational networks by increasing the array of monochromatic circles and using only the sameness cue, S, to relate particular circles to one another. For example, as in Carpentier et al., an array of nine circles could be divided into three, three-member sets; participants would refer to the relational networks to determine equivalence or non-equivalence between compound elements, followed by equivalence-equivalence or non-equivalence-non-equivalence testing and training.

In Experiment 1 of the present study, we sought to extend Study 3 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 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 replicated Experiment 1, but participants were children with autism spectrum disorder (ASD). Characterised by impairments in social interaction and social communication (American Psychiatric Association, 2013), ASD currently affects one 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 behavioural 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 generalised 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 contrast, in the present study, 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 pre-assessment testing, and of those, three passed the pre-assessment 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 Study 1 of this thesis. The researcher administered all probe and training sessions 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 three 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 the 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 consent was obtained from each participant.

Experimental Design

A multiple baseline design across participants was used in this study. Details for each condition of the multiple baseline are described below. 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 & Apparatus

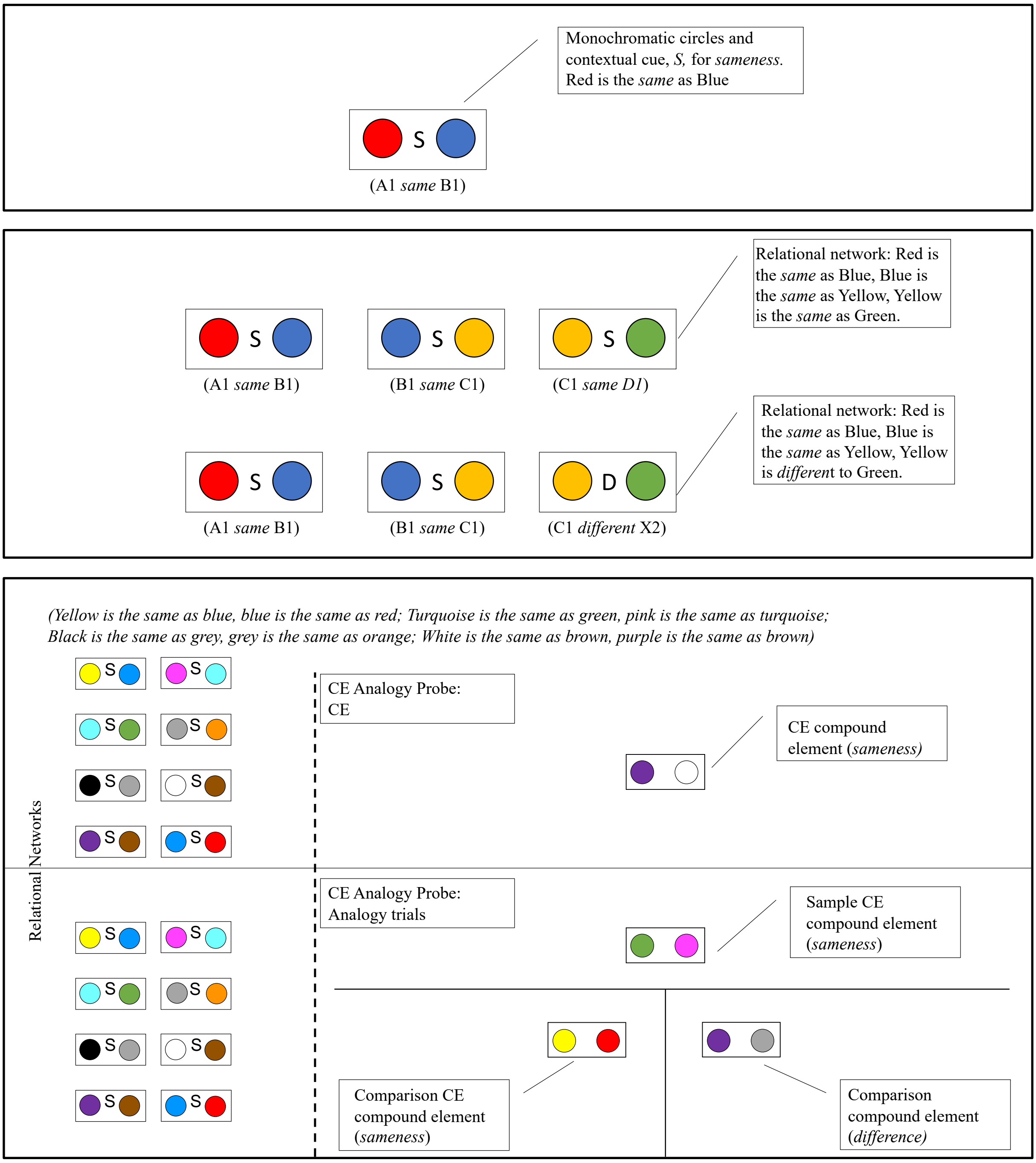
A 13-inch MacBook running Microsoft PowerPoint was used to present trials. Individual stimuli included coloured circles (either 0.5 inch or one inch in diameter, depending on the trial type) and letters 'S' for sameness and 'D' for difference relations in Calibri or Arial, size 24 or 36 fonts (depending on the trial type; see first panel in Figure 6.1). During the relational section of the pre-assessment, trials included an array of such stimuli at the centre of the screen that were designated by the experimenter as participating in a relational network (see second panel in Figure 6.1). During the analogical section of the pre-assessment, as well as during the 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 Figure 6.1 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] [Grey 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] [Grey Circle]. This array might allow a potential participant to derive four equivalence relations: black, grey, 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 Figure 6.1). 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.

Figure 6.1

Pre-Assessment Screening Tool Stimuli Arrangement



Procedure

Overview

The following will provide procedural details of the relational pre-assessment and the multiple baseline across participants design of the study. The pre-assessment was administered first in one session and took between 5 and 20 minutes to complete. Following pre-assessment, participants entered the baseline condition of the multiple baseline design, followed by a brief pre-training probe, followed by the intervention condition, which included training and probe sessions. Table 6.1 provides a schematic overview of procedures. An average of eight sets was run per day during the intervention condition; a probe set took on average three to four minutes to complete, and a training set took approximately one to two minutes to complete. Both participants started baseline sessions at the same time, and both participants took approximately three weeks to complete the study once the training condition was implemented.

Table 6.1

Schematic Presentation of Pre-Assessment 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

Baseline condition	
CE Probe Set 1: 6 CE trials + 6 CE analogy trials	100
Pre-Training DP Analogy Probe: 6 trials	100
Intervention	
DPA-Training Phase 1: 6 trials	100
DPA+XF Training Phase 2: 6 trials	100
Probes	
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

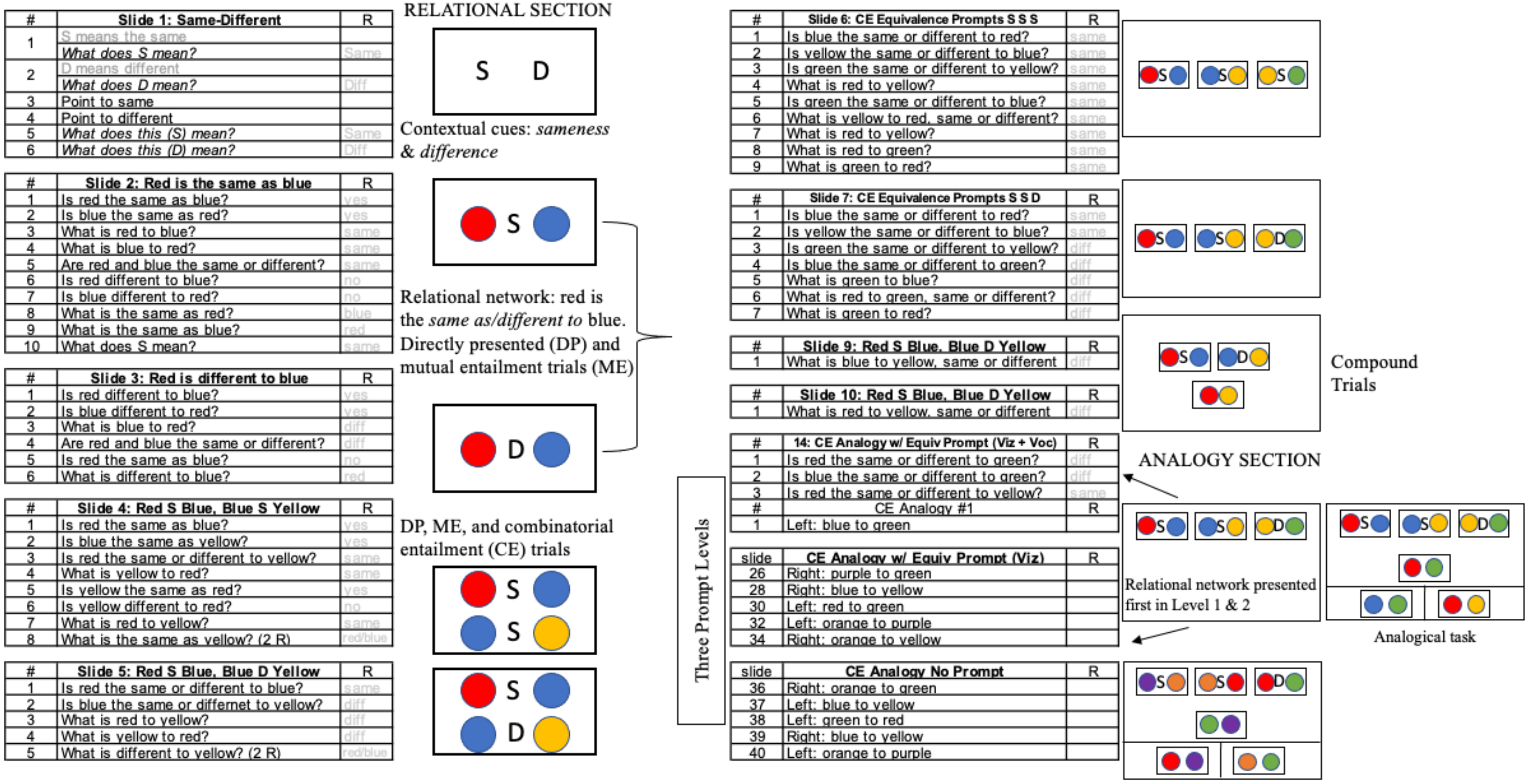
Note. DPA = directly presented relations; ME = mutually entailed relations; CE = combinatorially entailed relations.

Pre-Assessment

Tasks adapted from the arbitrary relational and analogical stages in Study 1 of this these were used in the relational pre-assessment (see Figure 6.2). The pre-assessment, which comprised 82 trials in total, included four main sections outlined in the following.

Figure 6.2

Pre-Assessment Screening Tool



RELATIONAL SECTION

Contextual cues: *sameness* & *difference*

Relational network: red is the *same as/different to* blue. Directly presented (DP) and mutual entailment trials (ME)

DP, ME, and combinatorial entailment (CE) trials

Compound Trials

ANALOGY SECTION

Three Prompt Levels

Relational network presented first in Level 1 & 2

Analogical task

#	Slide 1: Same-Different	R
1	S means the same	
	What does S mean?	Same
2	D means different	
	What does D mean?	Diff
3	Point to same	
4	Point to different	
5	*What does this (S) mean?*	Same
6	*What does this (D) mean?*	Diff

#	Slide 2: Red is the same as blue	R
1	Is red the same as blue?	yes
2	Is blue the same as red?	yes
3	What is red to blue?	same
4	What is blue to red?	same
5	Are red and blue the same or different?	same
6	Is red different to blue?	no
7	Is blue different to red?	no
8	What is the same as red?	blue
9	What is the same as blue?	red
10	What does S mean?	same

#	Slide 3: Red is different to blue	R
1	Is red different to blue?	yes
2	Is blue different to red?	yes
3	What is blue to red?	diff
4	Are red and blue the same or different?	diff
5	Is red the same as blue?	no
6	What is different to blue?	red

#	Slide 4: Red S Blue, Blue S Yellow	R
1	Is red the same as blue?	yes
2	Is blue the same as yellow?	yes
3	Is red the same or different to yellow?	same
4	What is yellow to red?	same
5	Is yellow the same as red?	yes
6	Is yellow different to red?	no
7	What is red to yellow?	same
8	What is the same as yellow? (2 R)	red/blue

#	Slide 5: Red S Blue, Blue D Yellow	R
1	Is red the same or different to blue?	same
2	Is blue the same or differnet to yellow?	diff
3	What is red to yellow?	diff
4	What is yellow to red?	diff
5	What is different to yellow? (2 R)	red/blue

#	Slide 6: CE Equivalence Prompts S S S	R
1	Is blue the same or different to red?	same
2	Is yellow the same or different to blue?	same
3	Is green the same or different to yellow?	same
4	What is red to yellow?	same
5	Is green the same or different to blue?	same
6	What is yellow to red, same or different?	same
7	What is red to yellow?	same
8	What is red to green?	same
9	What is green to red?	same

#	Slide 7: CE Equivalence Prompts S S D	R
1	Is blue the same or different to red?	same
2	Is yellow the same or different to blue?	same
3	Is green the same or different to yellow?	diff
4	Is blue the same or different to green?	diff
5	What is green to blue?	diff
6	What is red to green, same or different?	diff
7	What is green to red?	diff

#	Slide 9: Red S Blue, Blue D Yellow	R
1	What is blue to yellow, same or different	diff

#	Slide 10: Red S Blue, Blue D Yellow	R
1	What is red to yellow, same or different	diff

#	14: CE Analogy w/ Equiv Prompt (Viz + Voc)	R
1	Is red the same or different to green?	diff
2	Is blue the same or different to green?	diff
3	Is red the same or different to yellow?	same
#	CE Analogy #1	R
1	Left: blue to green	

slide	CE Analogy w/ Equiv Prompt (Viz)	R
26	Right: purple to green	
28	Right: blue to yellow	
30	Left: red to green	
32	Left: orange to purple	
34	Right: orange to yellow	

slide	CE Analogy No Prompt	R
36	Right: orange to green	
37	Left: blue to yellow	
38	Left: green to red	
39	Right: blue to yellow	
40	Left: orange to purple	

Note. Adapted from Study 1.

Section 1: Relational Tasks. The first six trials in the pre-assessment introduced the participant to the contextual cues themselves (i.e., S and D; refer to Figure 6.2). 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 Figure 6.2). 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 Figure 6.2). The first set of 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?’).

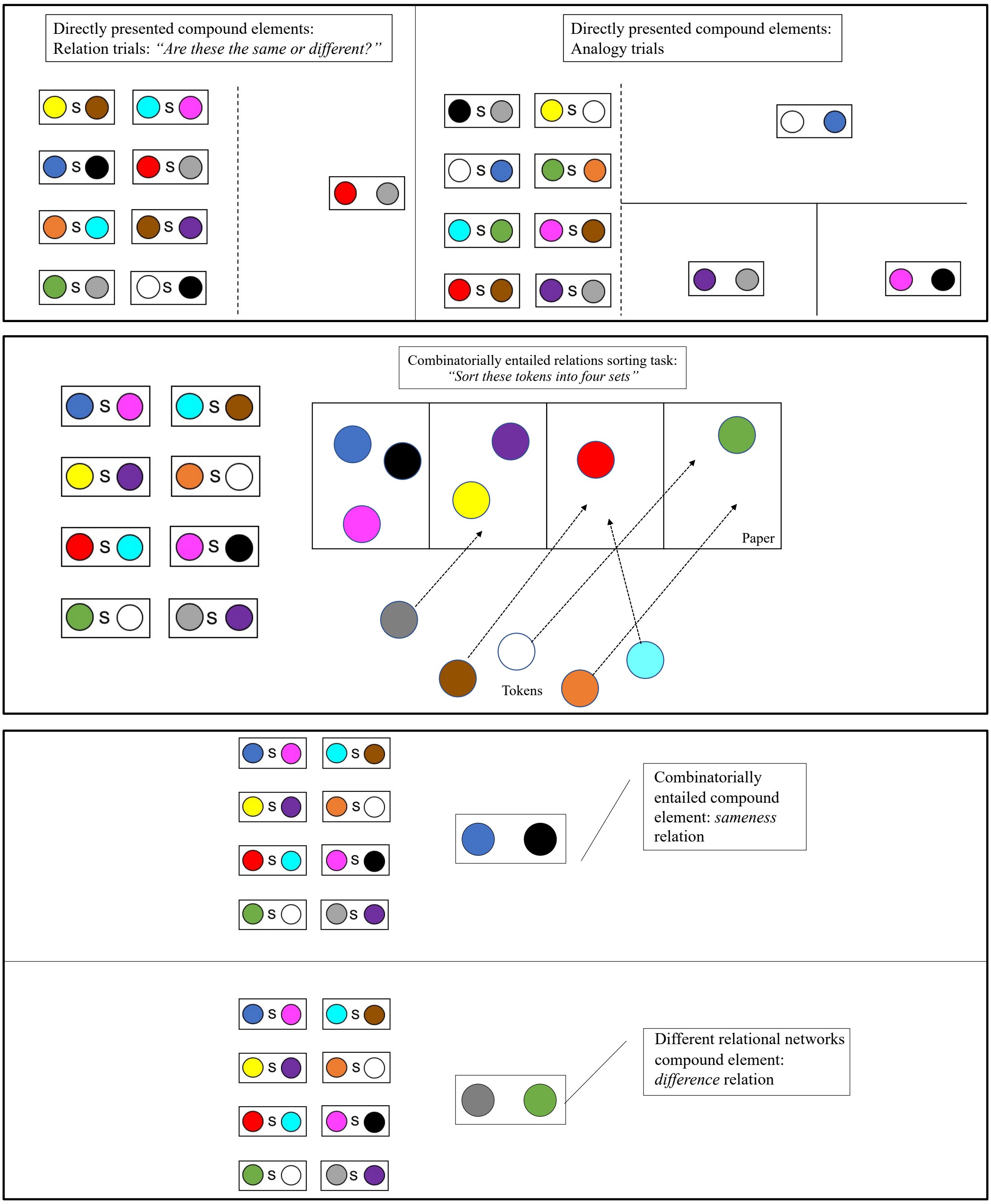
Section 2: Directly Presented Compound Elements. The next set of tasks presented compound elements, including a stimulus composed of two side-by-side monochromatic circles identical to the circles in the relational network, without the relational cue, S (see Section 2 in Figure 6.2). For example, one 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 coloured 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] [Grey 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] [Grey Circle] are presented at the left of the screen, and the compound stimulus (e.g., [Pink Circle][Grey Circle]) is presented to the right of the network; thus, the participant might look at the relational networks and find that pink and grey are the same (the compound element) because the relation is directly presented as [Purple Circle] [S] [Grey Circle]. Each slide included the four relational networks and one compound stimulus.

Following the compound questions, the second task in section two presented directly presented analogical stimuli (see Section 2 in Figure 6.2 and the first panel in Figure 6.3). The directly presented analogical stimuli were presented to the right of the relational networks. For example, the relational networks [Black Circle] [S] [Grey 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] [Grey 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][Grey 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 Figure 6.3). The participant had to refer to the relational networks to determine if the stimuli within the compound elements were the same or different.

Figure 6.3

Pre-Assessment Section 2: Directly Presented Analogy; Section 3: Combinatorial Entailment



Section 3: Combinatorial Entailment. In the first task in this section, participants were given 12 monochromatic tokens that matched the colours of the circles in the relational networks and a sheet of paper divided equally into four sections (see second panel of Figure 6.3). One token from each relational network was placed in its own section on the paper. The

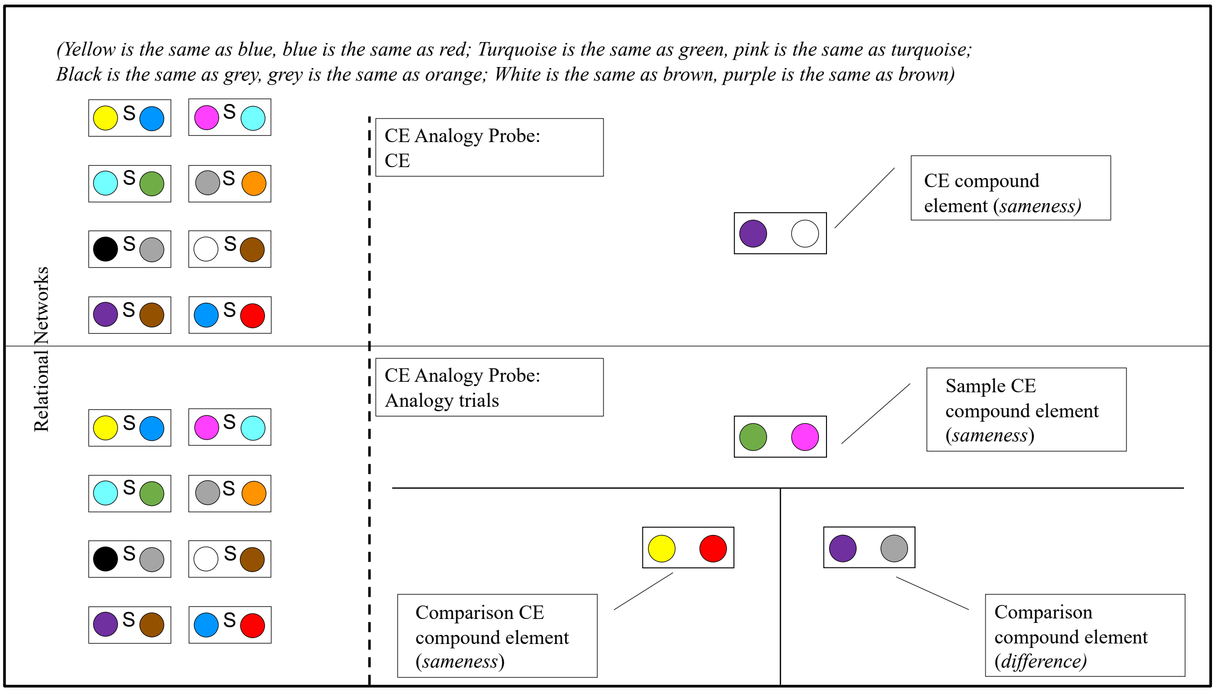
researcher gave the instruction, ‘Look here (points to the relational networks) to figure out which circles go with each other. There are four sets of circles and three circles in each set.’ Sorting responses were scored as correct or incorrect for a total of 12 responses.

The next task in Section 3 required the same sorting task followed by six questions regarding the combinatorially entailed relations among the stimuli. A PowerPoint 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 Figure 6.3).

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 re-administered. Potential participants had to score at least 80% correct to proceed to Section 4.

Figure 6.4

Pre-Assessment Section 4: Analogical Relations

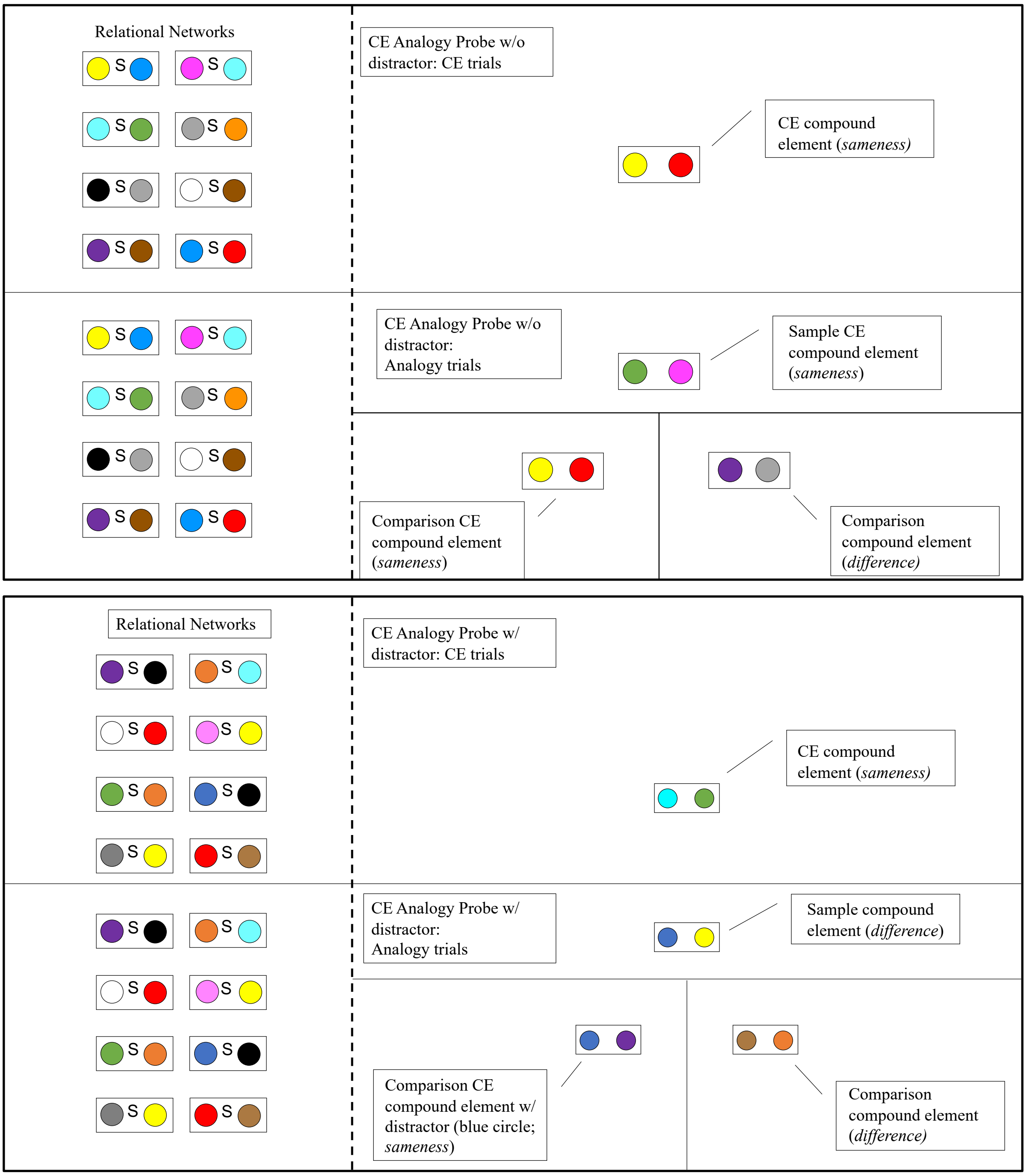


(Yellow is the same as blue, blue is the same as red; Turquoise is the same as green, pink is the same as turquoise; Black is the same as grey, grey is the same as orange; White is the same as brown, purple is the same as brown)

Section 4: Relating Combinatorially Entailed Relations: Analogical Relations. There were two tasks in the analogical section of the pre-assessment. 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 Figure 6.4). 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.' To proceed to the baseline condition , participants had to score at least five out of six correct (83%) on the CE relational trials and fail the analogy trials.

Figure 6.5

Two Probe Types: CE Analogy and CE+D Analogy



Multiple Baseline Conditions

The multiple baseline design across participants comprised a baseline condition including unreinforced baseline sessions, a brief, two-session pre-training 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 Figure 6.5 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 pre-assessment (Section 4 of the pre-assessment). 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 comprised 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 pre-assessment.

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 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 pre-training probe condition was implemented to assess participant responding to directly presented analogical relations (DPA-Probe). The stimulus format was the same as in the CE Probes except that all stimuli in the compound elements in the pre-training 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 pre-training probe sessions. The training condition included 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 Figure 6.6). 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 consequated with, ‘Yes, that is correct!’ and an incorrect trial was consequated with, ‘No, that is incorrect.’ The following trial was presented regardless of correct or incorrect responses.

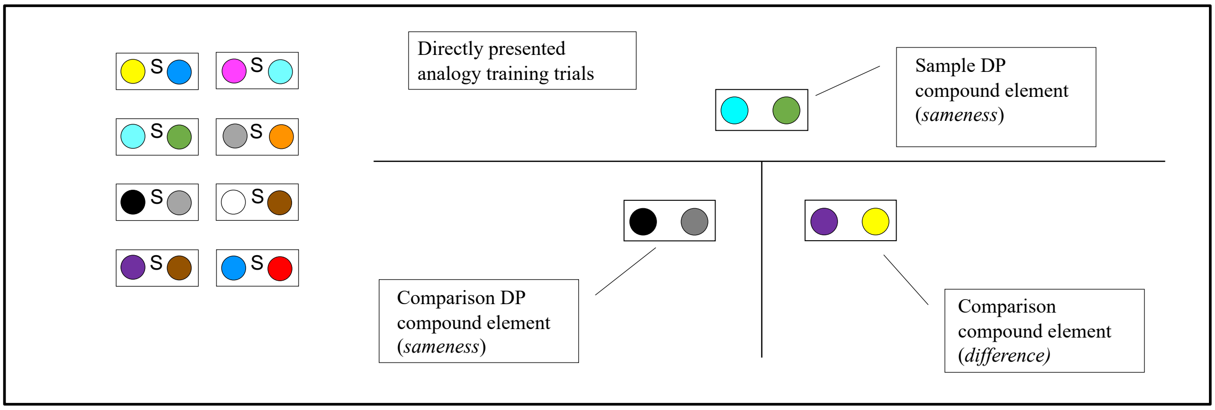
In Phase 2 DPA+XF Training, more instruction and feedback were included in each trial. The participant was shown the screen with the analogical stimuli and given the instruction, ‘Look at this one first (points to sample), and figure out if it’s the same or different. Now look at these here (points to 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 consequated 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).'

The passing criterion was 100% correct once on the DPA-Training trials. When the participant met the criteria, the baseline probe, CE Probe Set 1, was re-administered, including the six relational trials and the six analogy trials. If participants failed the CE Probe, they went back into training and had 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 pre-assessment for screening potential participants; a baseline condition in which relating combinatorially entailed relations (CE analogy) was tested; a brief pre-training 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.

Figure 6.6

Training Stimuli: Directly Presented Compound Elements for Directly Presented Analogical Responding



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 using 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%. Interobserver agreement was calculated on a trial-by-trial basis for each probe and training trial. Interobserver agreement was assessed for 28% of Participant P1.1's sessions, IOA was 100%; for 30% of P1.2's sessions, IOA was 100%.

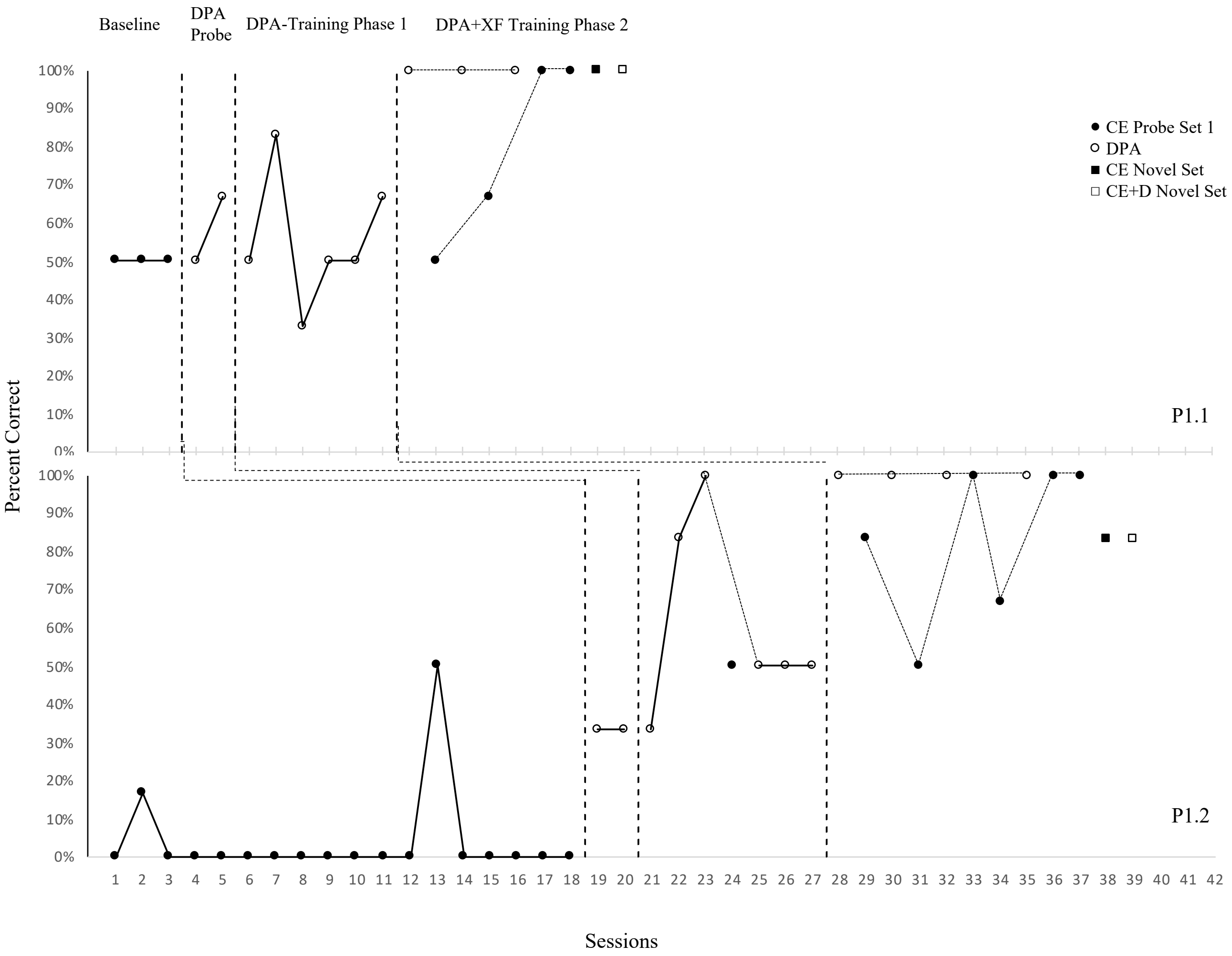
Results and Discussion

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 generalisation probe, CE+D Probe (see Table 6.2 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 generalisation probe, and Participant P1.2 scored 83% correct on the novel CE Probe and the generalisation probe. Both participants required Phase 2 DPA+XF Training (see Figure 6.7).

Figure 6.7

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

Table 6.2

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

Pre-Assessment

As previously indicated, six potential participants were tested for relational responding on the pre-assessment. Three of these children passed the combinatorially entailed relational tasks, and two of those proceeded to baseline sessions. Of the participants who met the criteria on the pre-assessment, 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 6.3 for pre-assessment scores.

Table 6.3

Experiment 1: Pre-Assessment Relational and Analogical Scores (Percent Correct)

Participant	Sort Tokens	CE Sort w/ Tokens	CE Sort w/o Tokens		CE & Analogy Probe	
		Relational	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.

Participant 1.1

Participant P1.1 scored 50% correct on all baseline combinatorially entailed analogical relations sessions (CE Probes) and 50% and 67% on the pre-training Directly Presented Analogy Probe (DPA-Probes) sessions (see Figure 8 for participant results). Participant P1.1 did not meet the training criteria during the Directly Presented Analogy Probe (DPA-Probes) pre-training 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 pre-training 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 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 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 generalisation data. Both participants passed baseline CE Probe Set 1 as well as a novel CE Probe following DPA analogy training. Furthermore, correct analogical responding generalised to the CE+D Probe.

This extends Studies 2 and 3 of this thesis in which in which an RFT approach was used to examine analogy in young children. Furthermore, these data support the Carpentier et al. (2002) and the Studies 2 and 3 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 behaviour analytic services. Participant P2.1 attended a private behavioural and learning centre in New York City, five days a week, for five hours 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 72nd percentile for 1st-grade reading and in the 27th percentile for 2nd-grade reading, and he scored in the 54th percentile for 3rd-grade math computation. Participant P2.2 scored in the 4th percentile for 3rd-grade reading and below the 1st percentile for 3rd-grade math computation. Participants were selected for inclusion based on their performance on an adaptation of the relational assessment used in Study 1 of this thesis. The researcher administered all probe and training sessions in an otherwise unoccupied room of Participant P2.1's centre, 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 three 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 the recruitment of participants was obtained from the director of the clinic, parental consent was obtained for each child who participated, and verbal consent was obtained from each participant.

Experimental Design

As in Experiment 1, a multiple baseline design across participants was used in Experiment 2. Experiment 2 included the same relational pre-assessment for screening potential participants; a baseline condition in which relating combinatorially entailed relations was tested; a brief pre-training 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 generalisation probe trials were presented. 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 & 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 pre-assessment was administered first in one session and took approximately 20 minutes to complete. Following pre-assessment, participants entered the baseline condition of the multiple baseline design, followed by a brief pre-training probe, followed by the intervention condition, which included training and probe sessions. Table 6.1 (Experiment 1) shows a schematic overview of procedures. An average of eight sets was run per day during the intervention condition; a probe set took on average three to four minutes to complete, and a training set took approximately one to two minutes to complete. Both participants started baseline sessions at the same time, and both participants took approximately one week to complete the study once the training condition was implemented (i.e., based on the administration of 4 – 8 probe and training sessions per day).

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 using 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%. Interobserver agreement was calculated on a trial-by-trial basis for each probe and training trial. Interobserver agreement was assessed for 48% of Participant P2.1’s sessions, IOA was 97%; for 21% of Participant P2.2’s sessions, IOA was 100%.

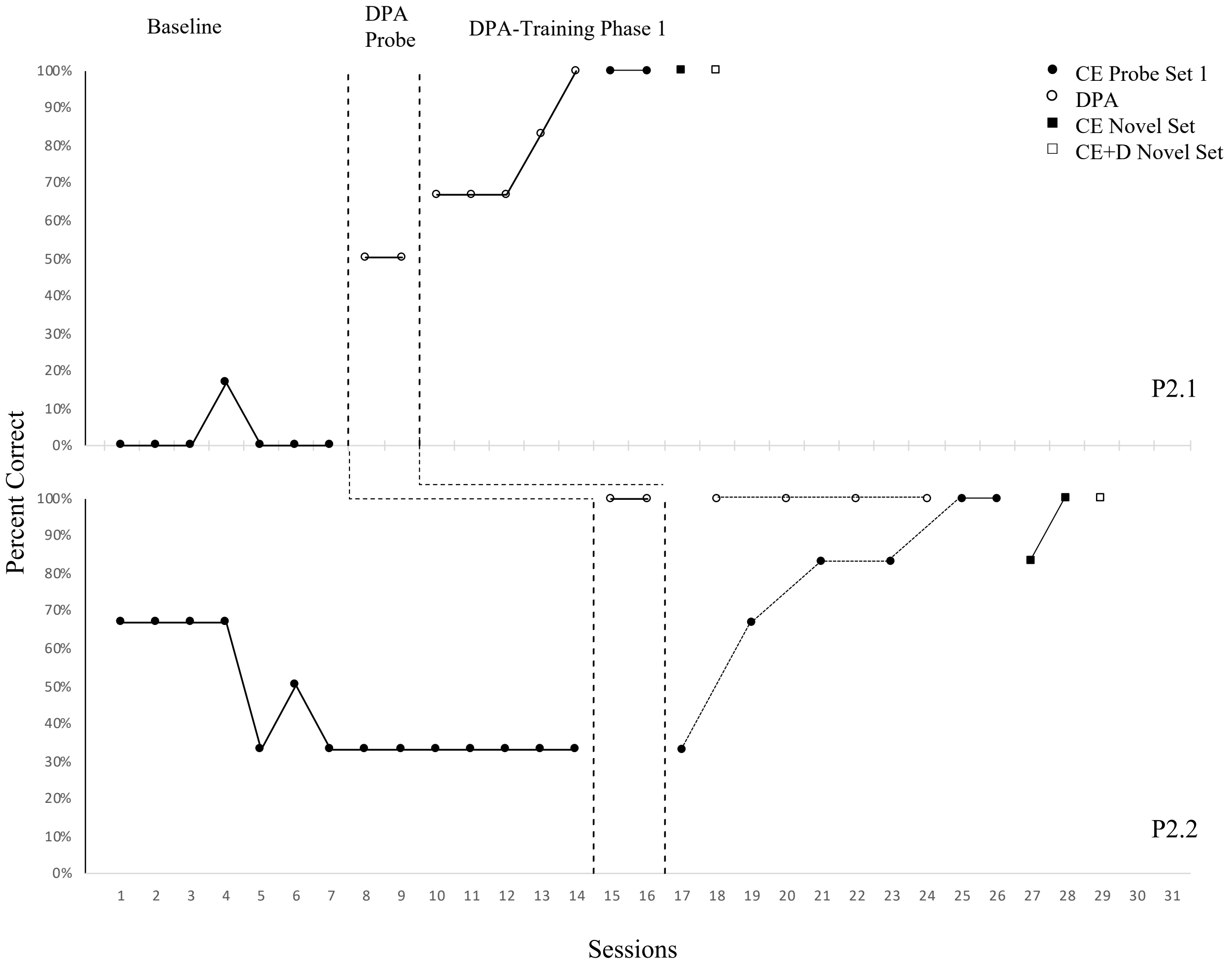
Results and Discussion

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 generalisation probe, Combinatorially Entailed Probe w/ Distractor (CE+D Probe; see Figure 6.8). 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 generalisation probe (refer to Table 6.2 for condition acronyms).

Figure 6.8

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

Pre-Assessment

Both participants were tested for relational responding on the pre-assessment. 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 6.4 shows pre-assessment scores.

Table 6.4

Experiment 2: Pre-Assessment Relational and Analogical Scores (Percent Correct)

Participant	Sort Tokens	CE Sort w/ Tokens	CE Sort w/o Tokens		CE & Analogy Probe	
		Relational	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.

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 on both pre-training DPA-Probe sessions. Participant P2.1 required five DPA-Training sessions before meeting the passing training 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 stayed at low levels of responding during baseline CE Probe sessions, and he scored 100% correct on both pre-training 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 generalisation 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 generalised to the CE+D Probe.

General Discussion of Study 4

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 non-equivalence-non-equivalence (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 pre-training 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 generalisation as well as of multiple exemplar training if required.

Studies 2 and 3 of this thesis 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 pre-training, and once established, it allowed testing and training of unlimited novel analogies. Studies 2 and 3 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, Studies 2 and 3 tested for derived relations between difference relations only.

The present study sought to extend the REP methodology used in Studies 2 and 3 but with some modifications. First, the relational networks included a larger array of relational stimuli than those presented in Studies 2 and 3, 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 generalisation probes.

Second, in the present study, unlike in Studies 2 and 3, 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 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 behaviour 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 generalisation 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 behavioural study of figurative language in children since Carpentier et al. is on metaphorical responding in children diagnosed with ASD (Persicke et al., 2012). However, the methodology used in Persicke et al. 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 these variables and thus allowed us to maintain experimental control while examining analogical responding in children with autism. Interestingly, the children with ASD in Experiment 2 required fewer DPA (i.e., 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 behaviour analytic interventions. Another possible contributing factor is the age difference of nearly 10 years between the typically developing children (approx. 5 years old) and the children with ASD (approx. 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 more generally, 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, generalisation 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 post-training sessions. It is possible that training children with ASD or other developmental delay using procedures such as the present one might similarly result in generalisation to the understanding and creation of novel figurative language in a more naturalistic context. This might be a focus of future work.

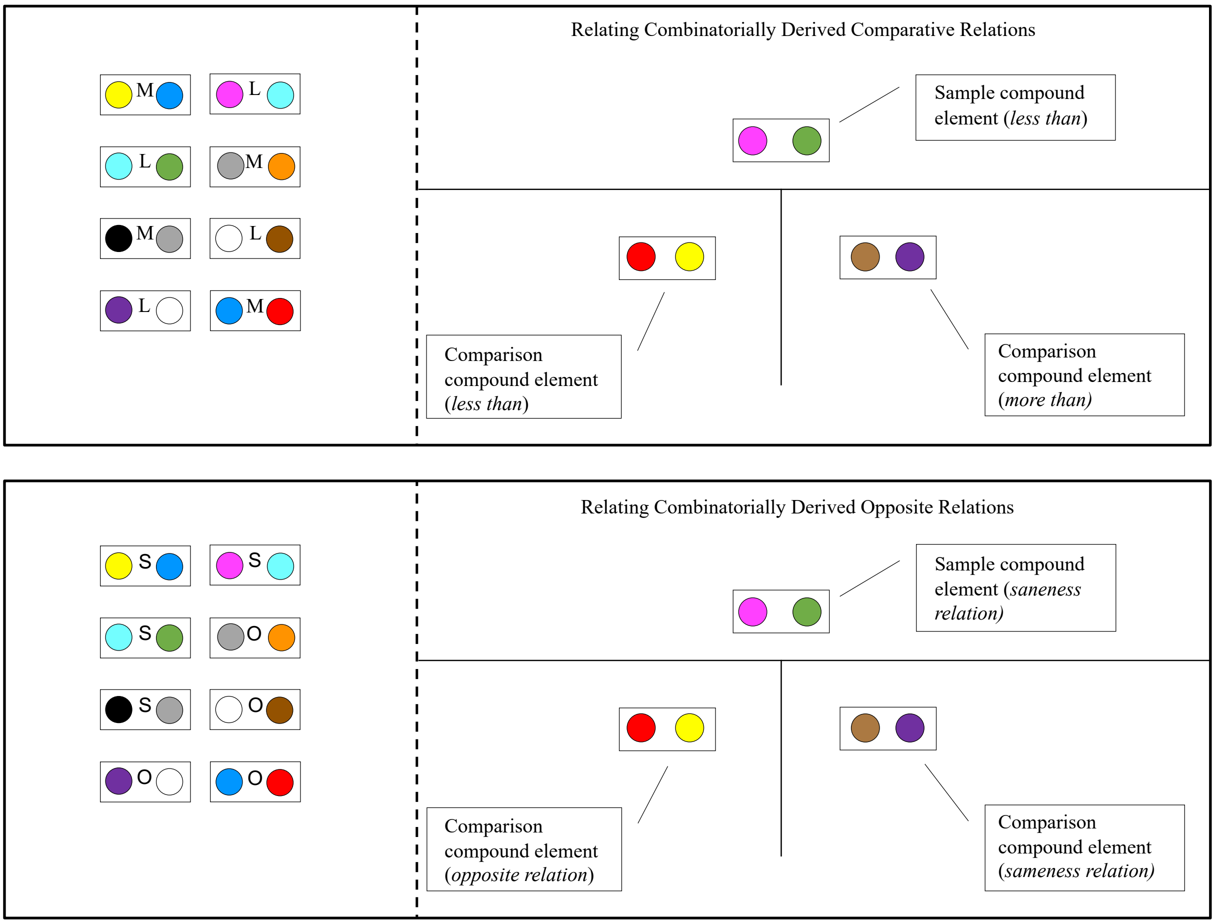
In both experiments of the present study, only Participant P2.2 in Experiment 2 did not require training on the directly presented analogy tasks as he scored 100% correct in both the DPA (directly presented analogy) probes and training trials. Following success on these trials, additional prompting with this task facilitated correct responding on the relation of combinatorially entailed relations. Participant P2.2 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, while the children failed initially to show the derivation of relations between derived relations without intervention, they did all spontaneously show the 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 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 generalised effects of training analogical responding on socially valid measures such as mainstream analogy tests, academic achievement tests, or standardised 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., Hayes &

Stewart, 2016; Cassidy et al., 2016) used the REP to assess and train derived relational responding and compared scores on pre- and post-training standardised 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 Figure 6.9. 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.

Figure 6.9

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).

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 generalisation across relations could occur. In Study 1 of the present thesis, we 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. Alternatively, despite the empirical findings of Study 1, it might also be investigated whether, under particular circumstances, training analogy involving non-coordinate 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 generalisation in novel relational varieties of analogy, or whether training additional relational varieties of analogy might be required to promote generalisation.

Future researchers investigating the acquisition of analogy might also usefully consider the dimensions along which analogical relational responding can vary as described in the multi-dimensional (and latterly hyper-dimensional) multi-level (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 analyse 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 to include a larger sample that might allow better insight into the emergence of analogical responding in young children. Furthermore, the pre-assessment 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 pre-assessment, 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 pre-assessment might warrant further investigation. In Section 3 of the pre-assessment, participants were required to sort tokens into sets based on the relational information provided in the relational network (see second panel of Figure 6.3). Participants were given 12 monochromatic tokens that matched the colours of the circles in the relational networks and a sheet of paper divided equally into four parts (see second panel of Figure 6.3). 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, coloured 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 pre-training 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 generalisation as well as of multiple exemplar training if required. On the other hand, it could be argued that, while 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 pre-assessment in the present study for example) 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 children's repertoire 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, while 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 behavioural 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 thesis 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.

Chapter 7: General Discussion

This thesis aimed to investigate the testing and training of analogical relations at the age of emergence in young children. Analogical reasoning is ubiquitous in our daily language and cognition. Many of our most celebrated technological and medical advances originated from analogical thinking. However, the countless analogies that navigate us through our everyday experiences are much humbler and fit so seamlessly into our internal and external vernacular that we hardly notice them (unless they catch you by surprise, such as when you are sitting on a ski lift at the beginning of a blizzard and your younger brother mentions that 'This is probably what the olden days looked like… everything is black and white, like in the pictures'). Despite the universality of analogy in language and cognition, there is very little behavioural research on 1) the age of emergence and 2) the assessment and training of analogical reasoning. Study 1 investigated a novel relational assessment protocol to test for the age at which analogy is acquired in young children. Studies 2–4 aimed to investigate an RFT-based procedure to test and train analogy in young children; in Experiment 2 of Study 4, the procedure was extended to test and train analogy in children with autism spectrum disorders.

Analogical reasoning has been chiefly the province of cognitive science in which analogical processes are theoretically described using metaphorical language. For example, the consensus among cognitive scientists is that analogical thinking can be deconstructed into several basic component processes, including 1) one or more relevant analogues to be stored in memory must be accessed; 2) a familiar analogue is to be mapped onto a target analogue; 3) analogical inferences are to be made from the mappings, allowing new knowledge to fill gaps in understanding; 4) the inferences are to be evaluated and possibly adapted to fit the unique requirements of the target; 5) new categories and schemas may be generated as a result of the analogical reasoning (Gentner & Holyoak, 1997; Holyoak et al., 2001; Vosniadou & Ortony, 1989). From a behavioural analytic perspective, these allegorical concepts may facilitate communication within the cognitive sciences. Still, this terminology is inadequate for a functional, scientific enquiry of analogical responding (the orthogonality of these epistemological stances is discussed later in the chapter) (Hayes et al., 1988).

The present thesis examined analogical responding using an inductive, behavioural analytic approach, specifically, relational frame theory. According to RFT, reasoning by analogy can be functionally operationalised as 'deriving relations between derived relations' (Barnes et al., 1997). The present thesis investigated both derived relations between derived relations as well as derived relations between nonderived relations. Of particular interest to this thesis was the research by Carpentier et al. (2002, 2003), who investigated the development of analogical responding in young children and provided evidence that analogical reasoning emerges around age five. Study 1 in the present research examined the development of relational frames, including analogy, in young children between three and seven years of age and found similar results. Using the Carpentier et al. (2002) study as a framework, Studies 2–4 further investigated the assessment and training of analogical relations in 5-year-old children. Experiment 2 of Study 4 replicated Experiment 1 (Study 4) but with participants diagnosed with autism spectrum disorders.

This chapter includes a brief overview of Studies 1–4, including a summary of the findings in general. Finally, theoretical considerations, limitations, and possible directions for further research conclude the discussion.

Overview of Studies

Study 1 (Chapter 3)

The primary aim of Study 1 was to focus on analogy, or the relating of relations, as one essential pattern of relational responding in young children, while a secondary aim was to measure the development of relational framing more generally in young children. Study 1 examined five frames including coordination, comparison, opposition, temporality, and hierarchy at four different levels of complexity, two of which focused on analogy specifically (Level 2, nonarbitrary relating of relations, and Level 4, arbitrarily applicable relating of relations) while the other two focused on basic relational framing at both the non-arbitrary (Level 1) and arbitrary (Level 3) levels. Participants were young children aged 3-7 and a relational evaluation procedure (REP)-based training and testing format was employed. Participants' relational performance across and within all levels and frames was correlated with their age and intellectual performance, as assessed on a standardised test of intellectual ability. The total score for arbitrary analogical responding showed a slightly stronger correlation with IQ performance than basic arbitrary relations (Stage 3). In comparison, basic arbitrary relations showed a slightly higher correlation with age compared to arbitrary analogical relations. These data provide further evidence that analogical relations are undoubtedly tied with intellectual potential. Arbitrary analogy scores revealed a marked difference between the 4- to 5-year-old cohort and the 5- to 6-year-old cohort. These findings 1) contribute to the extant RFT research suggesting a developmental divide in the acquisition of analogical ability at around five years of age (Carpentier et al., 2002, 2003) and 2) were used to inform participant selection in Studies 2–4.

Relational Frame Theory (RFT) views the operant acquisition of various patterns of relational framing (frames) as key to linguistic and cognitive development, and it has explored the emergence of a range of psychological phenomena (e.g., analogy, perspective-taking) in these terms. Furthermore, RFT has provided substantive evidence that derived relational responding underlies human language and cognition (Hayes et al., 2001a), including work showing that training this repertoire can boost intellectual skills. However, despite the evidence that relational framing is a core skill underlying cognitive ability, research on the sequence of frame acquisition and the normative development of relational responding remains limited. One potentially significant advance for RFT research is to obtain more detailed information on the normative development of relational framing in childhood. Hence this was the second aim of Study 1, which sought to measure relational responding of various types and at four levels of complexity in young children.

The normative data suggest, consistent with RFT, that the capacity to engage in relational responding and analogical responding (both overall and across different frames) are established and strengthened via ongoing exposure to the typical socio-verbal environment. Relational performance is correlated with measured IQ, which supports the RFT concept that relational framing is critical to language and cognition. Furthermore, the arbitrary stages are more highly correlated with IQ than the nonarbitrary stages. Analyses across the relational frames and levels revealed several patterns. Nonarbitrary and arbitrary coordination emerge first, and temporality emerges last; there is a difference between nonarbitrary and arbitrary relational responding for frames of opposition and hierarchy—opposition scores are lower in the nonarbitrary stages and higher in the arbitrary stages, and the hierarchy scores are higher in the nonarbitrary stages but lowest in the arbitrary stages. These normative data could help inform a developmentally sequenced relational curriculum.

Study 2 (Chapter 4)

Analogical responding, or relating relations, as operationally defined by relational frame theory, is pervasive in everyday language and cognition and is a critical component in learning. However, very little research on the assessment and training of analogical relations in young children exists in the behavioural literature. Study 2 assessed and trained analogical responding in young, typically developing children. Three 5-year-old children were assessed and trained in relating relations using an RFT-based REP protocol in a combination multiple baseline design across participants and a multiple probe design across behaviours. The study included a relational pre-assessment to screen potential participants; a baseline condition in which analogy was tested; and a training condition in which analogical responding was trained and generalisation probe trials (including three different probes; CE, DMC, and D-Cue Probes) were presented.

After training in relating combinatorially entailed relations, all three participants showed analogical responding according to RFT's conception of analogy as the derived relating of relations.

One crucial feature of Study 2 was that analogical responding was tested and trained using an adaptation of the novel RFT-based protocol used in Study 1. The utilisation of shapes across all trials obviated the need to control for participants' previous experience and knowledge and also controlled for stimulus consistency. The results of this study suggest the applied and experimental potential of this format. By using monochromatic circles and single-lettered contextual cues, the REP format makes it easy to produce multiple training and testing sets suitable for young children and pre-readers, making it possible to directly train children on complex relations with many exemplars. Moreover, this format does not require any relational pre-training, allowing for rapid testing (with multiple versions of tests). Finally, the REP format permitted a multiply controlled study in which we could target analogy testing and training directly while maintaining experimental control.

Study 3 (Chapter 5)

Study 3 was a replication of Study 2; however, the correction procedure was modified to provide more training opportunities for incorrect responding. In Study 2, the probe trials were presented whether the participant emitted a correct or incorrect response during the training correction procedure. In Study 3, the training procedure required participants to respond correctly before the probe was re-presented.

Three 5-year-old children were assessed and trained in relating relations using the RFT-based REP protocol in a combination multiple baseline design across participants and a multiple probe design across behaviours. As in Study 2, following multiple exemplar training, correct responding increased to criterion levels for all three children, and both generalisation and maintenance were observed.

The data from Studies 2 and 3 support the Carpentier et al. findings that 5-year-old children are capable of analogical responding. However, the present studies extend Carpentier et al. by directly training analogy through multiple exemplar training using a controlled multiple baseline design such that the participants of the study could thereafter demonstrate analogical responding without additional prompting procedures being needed.

It was noted in both Studies 2 and 3 that future work should consider increasing the number of stimuli in the relational networks to include more than four relata (i.e., the stimuli being related). Multiple relational networks could be presented by increasing the array of monochromatic circles and using only the sameness cue, S, to relate particular circles to one another. A larger array would allow for more combinatorially entailed sample and comparison compound combinations, thus circumventing a potential limitation noted in Studies 2 and 3. That is, all relational networks were composed of two sameness cues and one difference cue (e.g., Red S Blue, Blue S Green, Green D Yellow) and therefore permitted only combinatorially derived difference relations. Thus, the compound elements in the CE and D-Cue Probe trials had to be difference relations. Sameness relations were possible via directly presented, mutually entailed, or combinatorially entailed relations as in the DMC Probes. It is possible that participants learnt to pick the difference comparison compound instead of relating difference-difference relations. However, during the DMC Probes, it was possible for both sameness and difference relations to be tested because the stimuli in the sample and comparison compound elements included directly trained, mutually entailed, and combinatorially entailed relations. All participants passed the DMC Probes, which suggests that they were responding to analogical relations rather than just selecting the difference compound. However, in Studies 4 and 5, the number of relational networks was increased.

Study 4 (Chapter 6)

In Study 4, the testing and training procedure was modified to include a larger array of stimuli as suggested in Studies 2 and 3. The relational networks included 16 monochromatic circles and the relational cue, S for same, to delineate the relations between circles. A larger array would more closely replicate the analogical work by Carpentier et al. (2002), in which participants were trained and tested for equivalence-equivalence responding with nine arbitrary pictures. Furthermore, in Study 4, directly presented relations were trained, and derived relations were tested.

In Experiment 1 of Study 4, a multiple baseline design across participants was used in this study. The study included a relational pre-assessment for screening potential participants; a baseline condition in which relating combinatorially entailed relations (CE analogy) was tested; a brief pre-training 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 generalisation probe trials were presented. Two 5-year-old typically developing children were assessed and trained in relating relations in a multiple baseline design. 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 generalisation probe, CE+D Probe.

Experiment 2 was a replication of Experiment 1; however, Experiment 2 sought to investigate analogical responding in children diagnosed with autism spectrum disorders. Considering the prevalence of figurative language in our socio-verbal environment, children with ASD and other language delays face considerable comprehension challenges due to their difficulty with understanding non-literal language. Moreover, considering that analogy seems centrally important for language and cognition, training analogy in children struggling with language development could result in significant language generativity.

Two children with ASD were assessed and trained in relating relations. 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 generalisation probe, CE+D Probe. These results suggest that this format can be used to successfully train children with ASD to respond to analogical relations as defined by RFT. Future researchers could investigate whether training the core component processes of analogical responding in this way might facilitate training and understanding of more complex figurative language and whether analogical training results in the generative understanding of analogy and other figurative language in everyday language.

General Themes

This section explores the overarching themes that characterise this thesis, including the assessment of analogical relations in young children; the experimental and applied potential of the REP format for testing and training analogical relations in five-year-old children; this thesis' extension of previous behavioural work on analogical relations (i.e., Carpentier et al., 2002, 2003), and its contribution to the literature on analogical relations beyond behaviour analysis.

The present thesis contributes to the extant behavioural research work on analogical relations assessment and training in young children. The REP format used in this thesis permitted multiply controlled studies in which we could target analogy testing and training directly while maintaining experimental control. That is, the REP format afforded us with quick and effective stimulus control, allowing us to implement a multiple baseline design to examine the efficacy of multiple exemplar training to establish the core repertoire. The data from all five studies suggest the applied and experimental potential of this format.

Study 1 tested and confirmed the efficacy of the RFT-based REP format for assessing analogical relations, specifically and various patterns of relational framing more generally across increasingly complex levels. Data from the analogical stages support the findings of Carpentier et al. (2002) that analogical reasoning is acquired after age five. In addition, participant performance across frames and levels was analysed for 1) developmental sequences, 2) correlations with age and IQ, 3) intra-protocol correlations, and 4) performance across relational frames.

Studies 2–4 extended the Carpentier et al. (2002) research by successfully testing and training analogical relations in 5-year-old children using an adaptation of the REP-based format used in Study 1. These data provide further evidence that 5-year-old children are capable of analogical responding after training.

Previous RFT studies have successfully shown that many different patterns of relational framing can be tested and trained via multiple exemplar training (e.g., Belisle et al., 2016; Berens & Hayes, 2007; Cassidy et al., 2011, 2016; Mulhern et al., 2018). These include not just simple frames such as comparison and opposition (e.g., Berens & Hayes, 2007; Cassidy et al., 2011, 2016) but also relatively more complex patterns of framing such as hierarchy (Mulhern et al., 2018) and deixis (Belisle et al., 2016). The results from the present study have provided evidence that derived relating of relations can also be trained using multiple exemplar training.

The utilisation of shapes in the REP format in all five studies obviated the need to control for participants' prior history with language. The data from all five studies suggest the applied and experimental potential of this format. For example, by using monochromatic circles and single-lettered contextual cues, the REP format makes it easy to produce multiple training and testing sets suitable for young children and pre-readers, which, in turn, makes it possible to directly train children on complex relations with many exemplars. Moreover, this format does not require any relational pre-training but instead allows for faster testing (with multiple versions of tests) than is possible with the complex methodological issues in MTS procedures. And finally, a participant's previous experience and knowledge is not an experimental confound. These favourable variables of the REP format permitted multiply controlled studies in which we could target analogy testing and training directly while maintaining experimental control. That is, the REP format afforded us with quick and effective stimulus control, allowing us to implement a multiple baseline design to examine the efficacy of multiple exemplar training to establish the core repertoire. In addition, the REP format provided design flexibility for different generalisation probe types. Finally, as demonstrated in Study 1, this format can be applied to different relational frames at different levels of responding. These findings suggest considerable experimental and applied potential in language and cognition studies.

The present thesis contributes to the extant behavioural research work on relational language assessment and training. Carpentier et al. (2002, 2003) stripped analogical responding down to its simplest form and examined the basic relational processes required for analogical responding. The Carpentier et al. studies provided a systematic, bottom-up approach to studying analogy, but once you start getting into more complex versions of analogical responding, new questions arise. The definition they came up with is a result of the equivalence era they were in.

The present thesis looked at different variations of analogy, including different relations; however, even this is still just scratching the surface.

Previous RFT work on relating relations (e.g., Barnes et al., 1997; Carpentier et al., 2002) typically involved testing and training at least two sets of combinatorially entailed relations. The present work included directly presented, mutually entailed, and combinatorially entailed relations to provide evidence that participants were generalising analogical responding across different types of relational derivation rather than limiting their responses to the CE relation only. In Studies 2 and 3, training included combinatorially entailed relations between combinatorially entailed relations, and in Study 4, training included directly presented relations between directly presented relations. In all studies, all participants were successfully trained in analogical responding. These data suggest that regardless of the level of complexity of the derivation required (i.e., whether directly presented, mutually or combinatorially entailed), relations between relations should be considered analogy. Thus, derived relations between nonderived relations are also analogies, albeit simpler than derived relations between derived relations.

Barnes-Homes et al. (2017) include a discussion on the ‘derivedness’ of a derived relation. For example, once a relation has been derived the first time, is it still a derived relation? Furthermore, once it has been ‘derived’ many times over, what type of derived relation is it then? Future researchers investigating the acquisition of analogy might also usefully consider the dimensions along which analogical relational responding can vary as described in the multi-dimensional (and latterly hyper-dimensional) multi-level (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-practised’ 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 thesis, future researchers might refer to the MDML-HDML framework to experimentally analyse 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 concerning the training of other dimensions (e.g., relational flexibility).

Another consideration regarding RFT's definition of analogy is derived relations between derived relations. Is it a derived relation between a derived relation? It's not the same as other types of derived relations because you have two relations that are the same as each other—they are in a frame of coordination; this is more like a nonarbitrary relation—in other words, there is no cue that this is the same as that. Though at first glance, it might be tempting to say that two compounds are the same as each other because they contain the same relation, the reality is a little bit more nonarbitrary.

In daily language use, the relation between the elements in the source is usually known. There is no need for derived relation; thus, the RFT definition of analogy may consider extending to including direct and mutually entailed relations between direct and mutually entailed relations. In the present thesis, the REP procedure allowed for testing direct, mutually entailed, and combinatorially entailed relations. When you say the two relations are the same because they're both the same relation, that's not really the same as a derived relation.

The present data also constitute an important addition to the literature on analogical relations beyond behaviour analysis. Cognitive researchers have studied the acquisition of analogical responding in young children since Piaget's early work in the 1970s (e.g., Piaget et al., 1977/2001; Goswami, 1989, 1996, 2001; Rattermann & Gentner, 1998; Richland et al., 2006). However, as has been argued in previous RFT literature, this approach to analogical development has failed to lead to a functional analytic model of this behaviour that might lend itself to training this repertoire. Moreover, within the cognitive literature, the age at which analogical responding develops in young children remains a debate amongst researchers (Richland et al., 2006), and flaws in testing methodologies have hindered progress in testing and training analogy in this population. Within the cognitive literature, the development of analogical responding in young children has been theorised and examined, and the use of analogies to solve daily problems encountered in life or at school has been examined, but there is a shortage of experimental work on effective training of analogy in young children. In contrast, RFT's functional analytic definition of analogy has allowed a reliable methodology for training analogical reasoning in this population. However, metaphorically speaking, cognitive science and behavioural science are not the same game, and therefore, cannot and should not be pitted against each other or even compared.

The differences between cognitive and behavioural psychologists refer to their philosophical assumptions, scientific goals, and accepted explanations for the phenomenon in question (Hayes & Brownstein, 1986). Cognitive researchers believe that understanding underlying mental mechanisms is useful as they lead to new questions that increase the prediction and influence of behaviour. However, mental mechanisms are not directly observed, and thus, the presence or absence of mental processes may only ever be inferred based on overt behaviour. To make inferences, though, one needs objective evidence of how mental processes influence behaviour. Cognitive researchers are not able to provide such empirical evidence.

Testing predictions derived from mental process theories may lead to new functional knowledge about the conditions under which learning occurs. Contextual researchers can use this functional knowledge in order to refine more complex concepts and procedures.

Also, cognitive theories of analogical responding could influence applied work—closer interactions between functional and cognitive researchers at the applied level would likely facilitate the development of cognitive theories, creating a mutually beneficial situation.

Pepper (1942) argued four orthogonal epistemologies, or world hypotheses, including mechanism, formism, organicism, and contextualism, each with its own root metaphor and truth criterion. Relational frame theory subscribes to the contextualistic worldview; the root metaphor being the ongoing act in context and the truth criterion being successful working. Cognitive psychology aligns itself with the mechanistic worldview, the root metaphor being the machine, and the truth criterion being the correspondence between hypothesis and experimental findings. The orthogonality of these world hypotheses and their subscribing psychologies, in this case, the behavioural and cognitive psychology fields, may be illuminated by a simple comparative metaphor: A checkmate is not a royal flush—a win indeed, but two different games. We cannot intelligently compare the work and findings of the behavioural sciences with that of the cognitive sciences. The very purpose of behavioural science is practical: to operationally define behaviour and effect socially significant behavioural change through an inductive approach (i.e., on the battleground), whereas cognitivists operate at the level of metaphors themselves in complex thought experiments from a deductive approach (i.e., from a distance in their ivory towers; see Appendix H).

What then is the purpose of including a review of the cognitive research on analogy in a behavioural paper? From a pragmatic, applied perspective, familiarity with the cognitive literature may provide inspiration and guidance for future, socially relevant functional analytic research. For example, to date, the behavioural explorations of complex verbal behaviour, namely through RFT, are still in the early stages of development. Most of what we have done is quite simplistic, whereas cognitive psychologists have spent a lot of time discussing more complex analogies. Thus, behaviourists might find it helpful to refer to the cognitive literature in order to inform the functional experimental design of the same concepts/theories. Questions such as: Why aren't young children able to respond analogically? What are the component skills required for analogy? At what age can we train it, and what are the most efficacious methods for training? Ultimately, where are we going with this line of questioning? Behaviourists could translate the extant cognitive work into functional terms in order to meet the requirements of the contextualistic worldview and provide a practical, applied approach to testing and training language.

Future Directions

Considering the ubiquity of analogical responding in everyday life, more research regarding its development and training in young children both with and without language delays is merited. The present thesis adds to the limited behavioural research on analogical responding and suggestions for future directions are presented below. One particular focus for further research is the use of the REP format for testing and training analogy in young children. The use of the REP format allows for the testing and training of relating relations beyond coordination and distinction and furthermore permits extentions of the present research in multiple directions of study.

Considering the significant correlation between analogical responding and IQ found in Study 1, further work in identifying and training deficits in analogical reasoning was warranted and was thus the focus of Studies 2 – 4. In Studies 2 – 4, analogical responding was assessed and trained in 5-year-old children. Studies 2 – 4 add to the limited behavioural research on analogical responding in young children and contribute further evidence that 5-year-old children can be trained in analogical responding. Moreover, by offering a precise, functional analytic model of analogy, the present studies arguably shed additional light on this phenomenon beyond that provided by mainstream cognitive psychological work. This work provides further evidence that analogy tends to be acquired around the age of five but goes beyond previous work in showing how multiple exemplar training of the core pattern involved might accelerate this acquisition. Considering the ubiquity of analogical responding in everyday life, more research regarding its development and training in young children both with and without language delays is merited.

One notable feature of Studies 2 – 4 was that analogical responding was tested and trained using a novel RFT-based protocol, and more specifically, an adaptation of the RFT-based protocol used in Study 1. The utilisation of shapes across all trials obviated the need to control for participants' previous experience and knowledge as well as controlled for stimulus consistency across trials. The results of these studies suggest the applied and experimental potential of this format. For example, by using monochromatic circles and single-lettered contextual cues, the REP format makes it easy to produce multiple training and testing sets suitable for young children and pre-readers, making it possible to directly train children on complex relations with many exemplars. Moreover, this format does not require any relational pre-training but instead allows for faster testing (with multiple versions of tests) than is possible with the tricky methodological issues in MTS procedures. Finally, a participant's previous experience and knowledge is not an experimental confound. These favourable variables of the REP format permitted multiply controlled studies in which we could target analogy testing and training directly while maintaining experimental control. That is, the REP format afforded us with quick and effective stimulus control allowing us to implement a multiple baseline design to examine the efficacy of multiple exemplar training to establish the core repertoire. For example, once participants were trained on the format of the REP in the pre-assessments, we were able to immediately implement analogical baseline testing across all participants and then implement controlled training and testing conditions with novel sets of stimuli. In addition, the REP format provided design flexibility for different generalisation probe types.

Future work could extend the REP format to test and train analogical relations beyond coordination and distinction. Previously, Lipkens and Hayes (2009) successfully looked at multiple relations (i.e., sameness, difference, comparison, and opposition) in analogy in adult participants. Lipkens and Hayes tested across selection-based and topography-based tasks, including selecting the correct relata, selecting the correct relational cue, and producing the correct relata. However, the Lipkens and Hayes procedure required extensive pre-training as well as reading and writing skills. For young children, it might be possible to test and train multiple relations and topography-based responses using variations of the REP format used in the present study. For example, in Lipkens and Hayes' selecting relata task, the participant was given the sample, a relational cue, and two comparisons; participants had to select the correct comparison based on the relational cue; in the selecting the correct relation task, the participant had to select the correct relational cue given the sample and the comparison stimulus; and in the producing relata task, the participant had to produce the correct relata given the sample stimulus and relational cue. The REP format would support these various tasks and make them accessible to young children. For example, in producing the relata tasks, the response could include shading in a blank circle with the correct colour, and in selecting the relation tasks, the same label format (i.e., S or D) used in the current study could be applied and extended to other relations (see for example Study 1 in Chapter 3).

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 generalisation across relations could occur. In Study 1 of the present thesis, we 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. Alternatively, despite the empirical findings of Study 1, it might also be investigated whether, under particular circumstances, training analogy involving non-coordinate 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 generalisation in novel relational varieties of analogy, or whether training additional relational varieties of analogy might be required to promote generalisation.

Considering the relevance of analogy to intellectual potential, future researchers could investigate the generalised effects of training analogical responding on socially valid measures such as mainstream analogy tests, academic achievement tests, or standardised 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., Hayes & Stewart, 2016; Cassidy et al., 2016) used the REP to assess and train derived relational responding and compared scores on pre- and post-training standardised 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 Figure 6.9. 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 subsequently examine the effects of such training on intellectual potential.

Experiment 2 of Study 4 extended the work on analogy to include children diagnosed with ASD. The closest behavioural study of figurative language in children since Carpentier et al. is on metaphorical responding in children diagnosed with ASD (Persicke et al., 2012). However, the methodology used in Persicke et al. 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 these variables and thus allowed us to maintain experimental control while examining analogical responding in children with autism. The results from Experiment 2 (Study 4) indicate that the REP 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 more generally, training analogy in children with deficits in language development could result in generativity and creativity in language skills in addition to encouraging intellectual growth. Future work could examine if training children with ASD or other developmental delays using procedures such as the present one might result in generalisation to the understanding and creation of novel figurative language in a more naturalistic context.

Future researchers investigating the acquisition of analogy might also usefully consider the dimensions along which analogical relational responding can vary as described in the multi-dimensional (and latterly hyper-dimensional) multi-level (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-practised’ 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 analyse 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).

Considering the potential for improving language and cognition, further research into relational assessment and training, in general, is undoubtedly warranted. Up until recently, most empirical work on verbal behaviour has primarily been influenced by Skinner’s (1957) analysis of verbal behaviour (Dymond et al., 2010; Dixon, Belisle, et al., 2017). Consequently, commonly used language assessments, such as the Assessment of Basic Language and Learning Skills – Revised (ABLLS-R; Partington, 2008) and the Verbal Behaviour Milestones and Placement Program (VB-MAPP; Sundberg, 2008), are based on Skinner's analysis of verbal behaviour; thus, training focuses on the basic verbal operants (i.e., echoics, mands, tacts, intraverbals) with little attention to more complex verbal behaviour. More recently, Dixon et al. (2014, 2018) and Dixon, Rowsey, et al. (2017) provided experimental work assessing and training more complex language. The present study contributes to the work on relational language assessment.

Previous relational training studies (e.g., Barnes-Holmes, Barnes-Holmes, & Smeets, 2004; Belisle et al., 2016; Berens & Hayes, 2007) showed that derived relational responding could be brought under operant control. A functional analysis of young or developmentally delayed children's existing relational abilities would provide the framework for a robust, flexible, and individualised RFT-based curriculum. The present study is one of the first studies to examine the sequence of acquisition of multiple frames in young children; these findings could be a valuable reference in curriculum design.

Relational frame theory proposes that fluent and flexible derived relational responding may underlie much of human cognition (Hayes et al., 2001); thus, a functional analysis of existing relational abilities would facilitate designing effective relational programmes. Furthermore, several studies have investigated the effects of training relational responding and have found that training in relatively few frames results in improvements across a wide range of intellectual and language processes. For example, Cassidy et al. (2011, 2016) and Hayes and Stewart (2016) found significant increases in IQ scores and intellectual performance for all participants after completing automated relational training. In addition, Mulhern et al. (2018) trained hierarchical framing in young children and saw increased scores in assessments of language and categorisation. Therefore, by testing a child’s relational abilities and understanding the normative development of relational frames, we can design a curriculum that targets specific relational frames and builds a functional relational repertoire without omitting the necessary component relational skills. Furthermore, by providing a general overview of a child’s relational repertoire, the assessment in this study could be the basis for an RFT-based curriculum for children who present with learning difficulties or with developmental and intellectual disabilities.

Conclusion

The present thesis adds to the limited behavioural research on analogical responding in young children and contributes further evidence that 5-year-old children can be tested and trained in analogical responding. The ubiquity of analogical responding in everyday life warrants further research regarding its development and training in young children with and without language delays. Furthermore, the potential of the REP format used in this thesis to test and train young children in complex relational responding, such as analogy, is promising and lends itself to further investigation of its research and applied utility. Study 1 contributes to the exiguous literature on relational testing and training and initiates RFT-based research on the normative development of frames. Research in RFT has produced considerable evidence that relational framing is a core skill underlying language and cognition. However, research into the assessment and training of relational responding remains limited. The results from Study 1 could be referred to as a starting point for further research in the normative development and sequence of relational frames and for relational skills assessment and training.

Finally, research procedures do not determine application but rather the interest shown by society in the problems being studied. Regarding behavioural application, the importance and benefits of applied RFT to children and society cannot be denied. The testing and training protocol used in the present thesis is successfully being used to design and create applied research protocols across relational frames including coordination, distinction, comparison, opposition, simile, categorisation and hierarchy, and deixies, from nonarbitrary levels to increasinly complex arbitrary levels.

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

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. *The Analysis of Verbal Behavior, 14*(1), 57–83. https://doi.org/10.1007/BF03392916

Barnes-Holmes, D., Barnes-Holmes, Y., Luciano, C., & McEnteggart, 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*(4), 434–445. http://dx.doi.org/10.1016/j.jcbs.2017.08.001

Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2020). Updating RFT (more field than frame) and its implications for process-based therapy. *The Psychological Record.* Advance online publication. https://doi.org/10.1007/s40732-019-00372-3

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 a generalized operant behavior in young children. *International Journal of Psychology and Psychological Therapy, 4*, 531–558.

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 Publishers.

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/jeab.2005.79-04

Bartha, P. (2019). Analogy and analogical reasoning. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2019 ed.). Stanford University. 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*(4), 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., Colbert, D., Stewart, I., & Grey, I. M. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learning and Individual Differences, 47,* 222–235. http://dx.doi.org/10.1016/j.lindif.2016.03.001

Cassidy, S., Roche, B., & Hayes, S. C. (2011). A relational frame training intervention to raise intelligence quotients: A pilot study. *The Psychological Record, 61*(2), 173–198.

https://doi.org/10.1007/BF03395755

Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science, 38*(2), 383–397. https://doi.org.10.1111/cogs.12099

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.

Dancygier, B., & Sweetser, E. (2014). *Cambridge textbooks in linguistics: Figurative language.* Cambridge University Press.
https://doi.org/10.1007/s10864-018-9302-2

Dixon, M. R., Belisle, J., & Stanley, C. R. (2018). Derived relational responding and intelligence: Assessing the relationship between the PEAK-E pre-assessment and IQ with individuals with autism and related disabilities. *The Psychological Record, 68*(4), 419–430. https://doi.org/10.1007/s40732-018-0284-1

Dixon, M. R., Belisle, J., McKeel, A., Whiting, S., Speelman, R., Daar, J. H., & Rowsey, K. (2017). An internal and critical review of the PEAK relational training system for children with autism and related intellectual disabilities: 2014-2017. *The Behavior Analyst, 40*(2), 493–521. https://doi.org/10.1007/s40614-017-0119-4

Dixon, M. R., Belisle, J., Whiting, S. W., & Rowsey, K. E. (2014). Normative sample of the PEAK relational training system: Direct training module and subsequent comparisons to individuals with autism. *Research in Autism Spectrum Disorders, 8*(11), 1597–1606. https://doi.org/10.1016/j.rasd.2014.07.020

Dixon, M. R., Rowsey, K. E., Gunnarsson, K. F., Belisle, J., Stanley, C. R., & Daar, J. H. (2017). Normative sample of the PEAK relational training system: Generalization module with comparison to individuals with autism. *Journal of Behavioral Education, 26*(1), 101–122. https://doi.org/10.1007/s10864-016-9261-4

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. http://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., May, R. J., Munnelly, A., & Hoon, A. E. (2010). Evaluating the evidence base for relational frame theory: A citation analysis. *The Behavior Analyst, 33*(1), 97–117. http://doi.org/10.1007/bf03392206

Dymond, S., & Roche, B. (Eds.). (2013). *Advances in relational frame theory*: *Research and application.* Context Press; New Harbinger Publications, Inc.

Engelmann, S., Becker, W. C., Carnine, D., & Gersten, R. (1988). The direct instruction follow through model: Design and outcomes. *Education and Treatment of Children, 11*(4), 303–317.

Fagot, J., & Maugard, A. (2013). Analogical reasoning in baboons (papio papio): Flexible reencoding of the source relation depending on the target relation. *Learning and Behavior, 41*, 229–237. https://doi.org/10.3758/s13420-012-0101-7

Fienup, D. M., & Brodsky, J. (2020). Equivalence-based instruction: Designing instruction using stimulus equivalence. In M. Fryling, R. A. Rehfeldt, J. Tarbox & L. J. Hayes (Eds.), *Applied behavior analysis of language & cognition* (pp. 157–173). Context Press; New Harbinger Publications, Inc.

Fryling, M., Rehfeldt, R. A., Tarbox, J., & Hayes, L. J. (Eds.). (2020). *Applied behavior analysis of language and cognition.* Context Press; New Harbinger Publications, Inc.

Garred, M., & Gilmore, L. (2009). To WPPSI or to Binet, that is the question: A comparison of the WPPSI-III and SB5 with typically developing preschoolers. *Australian Journal of Guidance & Counselling, 19*(2), 104–115. https://dx.doi.org/10.1375/ajgc.19.2.104

Gaynes, R. (2017). The discovery of penicillin: New Insights after more than 75 years of clinical use. *Emerging Infectious Diseases, 23*, 849–853.

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

Gentner, D., & Christie, S. (2010). Mutual bootstrapping between language and analogical processing. *Language and Cognition, 2*(2), 261–283. https://doi.org/10.1515/LANGCOG.2010.011

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

Gentner, D., Holyoak, K. J., & Kokinov, B. (Eds.). (2001). *The analogical mind: Perspectives from cognitive science.* MIT Press.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*(3), 45–56.

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 and 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. http://doi.org/10.1037/0097-7403.7.1.1

Gore, N. J., Barnes-Holmes, Y., & Murphy, G. (2010). The relationship between intellectual functioning and relational perspective-taking. *International Journal of Psychology and Psychological Therapy, 10*(1), 1–17.

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. http://dx.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*(1), 69–95. http://dx.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*(3), 207–226. http://dx.doi.org/10.1016/0010-0277(90)90057-Q

Green, A. E., Kramer, D. J., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2012). Neural correlates of creativity in analogical reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(2), 264–272. https://doi.org/10.1037/a0025764

Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding new verbal capabilities in children with language delays.* Pearson Education, Inc.

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/10.1111/bjep.12114

Hayes, S. C. (2001). A personal prologue. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 141–154). Kluwer Academic; Plenum Publishers.

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 Publishers.

Hayes, S. C., Gifford, E. V., Wilson, K. G., Barnes-Holmes, D., & Healy, O. (2001). Derived relational responding as learned behavior. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 21–49). Kluwer Academic; Plenum Publishers.

Hesse, M. (1965). Aristotle's Logic of Analogy. *The Philosophical Quarterly, 15*(61), 328–340. https://doi.org/10.2307/2218258

Hesse, M. (1996). *Models and analogies in science.* Notre Dame University Press.

Hey, J., Linsey, J., Agogino, A. M., & Wood., K. L. (2008). Analogies and metaphors in creative design. *International Journal of Engineering Education, 24*, 283–294.

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.

Holyoak, K. J. (2005). Analogy. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 117–142). Cambridge University Press.

Holyoak, K. J., Gentner, D., & Kokinov, B. N. (2001). Introduction: The place of analogy in cognition. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 1–19). The MIT Press.

Holyoak, K. J., & Thagard, P. R. (1989). A computational model of analogical problem solving. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning.* (pp. 242–266). Cambridge University Press.

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

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

Hughes, S., & Barnes-Holmes, D. (2016). Relational frame theory: The basic account. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes, & A. Biglan (Eds.), *The Wiley handbook of contextual behavioral science* (pp. 129–178). Wiley Blackwell.

Hunter, I. M. L. (1957). The solving of three term series problems. *British Journal of Psychology, 48*(4), 286–298. https://doi.org/10.1111/j.2044-8295.1957.tb00627.x

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. DOI: 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. http://doi.org/10.2307/1127862

Lindsley, O. R. (1992). Precision teaching: Discoveries and effects. *Journal of Applied Behavior Analysis, 25*(1), 51–57.

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

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., Rodriguez, M., Manas, I., Ruiz, F., Berens, N. M., & Valdivia-Salas, S. (2009). Acquiring the earliest relational operants: Coordination, distinction, opposition, comparison, and hierarchy. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding applications for learners with autism and other developmental disabilities* (pp. 149–172). Context Press; New Harbinger Publications, Inc.

Lunzer, E. A. (1965). Problems of formal reasoning in test situations. *Monographs of the Society for Research in Child Development, 30*(2), 19–46. http://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. 10.15585/mmwr.ss6904a

Matos, M. A., & Passos, M. L. P. (2010). Emergent verbal behavior and analogy: Skinnerian and linguistic approaches. *The Behavior Analyst, 3*3, 65–81. https://doi.org/10.1007/BF03392204

McCormack, T., & Hoerl, C. (2017). The development of temporal concepts: Learning to locate events in time. *Timing & Time Perception, 5*(3–4), 297–327. https://doi:10.1163/22134468-00002094

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 and Psychological Therapy, 4*, 303–323.

Meyer, C. S., Cordeiro, M. C., & Miguel, C. F. (2019). The effects of listener training on the development of analogical reasoning. *Journal of the Experimental Analysis of Behavior, 112*, 144–166. https://doi.org/10.1002/jeab.549

Moran, L., Stewart, I., McElwee, J., & Ming, S. (2014). Relational ability and language performance in children with autism spectrum disorders and typically developing children: A further test of the TARPA protocol. *The Psychological Record, 64*, 233–251. https://doi.org/10.1007/s40732-014-0032-0

Morsanyi, K., & Holyoak, K. J. (2010). Analogical reasoning ability in autistic and typically developing children. *Developmental Science, 13*(4), 578–587. https://doi.org/1-.1111/j.1467-7687.2009.00915.x

Mulhern, T., Stewart, I., & McElwee, J. (2017). Investigating relational framing of categorization in young children. *The Psychological Record, 67*, 519–536. https://doi.org/10.1007/s40732-017-0255-y

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'Hora, 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. http://doi.org/10.1007/BF03395504

O'Hora, D., Pelaez, M., Barnes-Holmes, D., Raw, G., Robinson, K., & Chaudhary, T. (2008). Temporal relations and intelligence: correlating relational performance with performance on the WAIS-III. *The Psychological Record*, *58*, 569–584.

Oppenheimer, J. R. (1956). Analogy in science. *American Psychologist, 11*, 127–135.

O'Toole, C., & Barnes-Holmes, D. (2009). Three chronometric indices of relational responding as predictors of performance on a brief intelligence test: The importance of relational flexibility. *The Psychological Record, 59*, 119–132.https://doi.org/10.1007/BF03395652

Partington, J. W. (2008). *The assessment of basic language and learning skills-revised (the ABLLS-R).* Behavior Analysts, Inc.

Pennypacker, H. S., Gutierrez, A., & Lindsley, O. R. (2003). *Handbook of the Standard Celeration Chart.* Cambridge Center for Behavioral Studies.

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.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)

Pyykkönen, P., & Järvikivi, J. (2012). Children and situation models of multiple events. *Developmental Psychology, 48*(2), 521–529. https://doi.org/10.1037/a0025526

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 Publications, Inc.

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

Richland, L., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*, 177–192. https://doi.org/10.1002/wcs.1336

Roid, G. H. (2003). *Stanford-Binet intelligence scales* (5th ed.). Riverside Publishing.

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 and Hearing Research, 14*(1), 5–13. https://doi.org/10.1044/jshr.1401.05

Skinner, B. F. (1957). *Verbal behavior*. Prentice Hall, Inc.

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

Spellman, B. A., & Schauer, F. (2005). Legal reasoning. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 719–735). Cambridge University Press.

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., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly, 47*(1), 1–41.

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

Stevenson, C. E., Alberto, R. A., van den Bloom, M. A., & de Boeck, A. L. (2014). Visual relations children find easy and difficult to process in figural analogies. *Frontiers in Psychology, 5*, 1–9. https://doi.org/10.3389/fpsyg.2014.00827

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., & Barnes-Holmes, D. (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). Kluwer Academic/Plenum Publishers.

Stewart, I., Barnes-Holmes, D., & Roche, B. (2004). A functional-analytic model of analogy using the relational evaluation procedure. *The Psychological Record, 54*(4), 531–552. https://doi.org/10.1007/BF03395491

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., & 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 Publications, Inc.

Stewart, I., McElwee, J., & Ming, S. (2013). Language generativity, response generalization, and derived relational responding. *The 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, M. Fryling, & L. J. Hayes (Eds.), *Applied behavior analysis of language and cognition: Core concepts & principles for practitioners.* Context Press; New Harbinger Publications, Inc.

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 Publications, Inc.

Sundberg, M. L. (2008). *VB-MAPP: Verbal behavior milestones assessment and placement program.* AVB Press.

Thoma, P., & Daum, I. (2006). Neurocognitive mechanisms of figurative language processing – Evidence from clinical dysfunctions. *Neuroscience and Biobehavioral Reviews, 30*, 1182 - 1205. doi:10.1016/j.neubiorev.2006.09.001

Vitale, A., Barnes-Holmes, Y., Barnes-Holmes, D., & Campbell, C. (2008). Facilitating responding in accordance with the relational frame of comparison: Systematic empirical analyses. *The Psychological Record, 58*, 365–390. https://doi.org/10.1007/BF03395624

Vosniadou, S., & Ortony, A. (Eds.). (1989). *Similarity and analogical reasoning.* Cambridge University Press.

Zettle, R. D., Hayes, S. C., Barnes-Holmes, D., & Biglan, A. (Eds.). (2016). *The Wiley handbook of contextual behavioral science.* Wiley Blackwell.

Appendices

Appendix A1

Relational Assessment Stage 1: Nonarbitrary Relations

Coordination

Example: Which one is like the one at the top (instructor points to sample)?

Comparison

Example: Which one is bigger/smaller than the one at the top (instructor points to sample)?

Opposition

Example: Which one is opposite to the one at the top (instructor points to sample)?

Temporality

Example: Which x is before/after the other y? Was the x before/after the y?

Hierarchy

Example: Which one is inside? Which one contains the other one? Does x contain y? Is x inside y?

Appendix A2

Relational Assessment Stage 2: Nonarbitrary Analogical Relations

Coordination

Example: Which one of these (points to comparisons) is like the one at the top (points to sample)?

Comparison

Example: Which one of these (points to comparisons) is like the one at the top (points to sample)?

Opposition

Example: Which one of these (points to comparisons) is like the one at the top (points to sample)?

Temporality

Example: Which one of these (points to comparisons) is like the one at the top (points to sample)?

Hierarchy

Example: Which one of these (points to comparisons) is like the one at the top (points to sample)

Appendix A3
Relational Assessment Stage 3: Arbitrary Relations

Coordination

Example: Does x like the same food as y? Does x like different food than y?

Comparison

Example: Which coin would you take to the store?

Opposition

Example: Which coin would you take to the store?

Temporality

Example: Is x before/after y?

Hierarchy

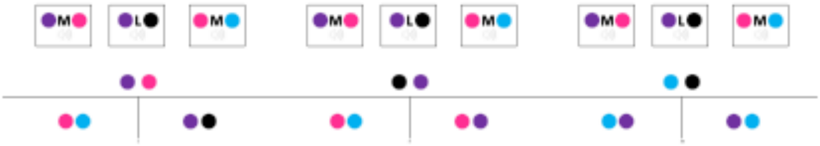
Example: Is x inside y? Does x contain y?

Appendix A4
Relational Assessment Stage 4: Arbitrary Analogical Relations

Coordination

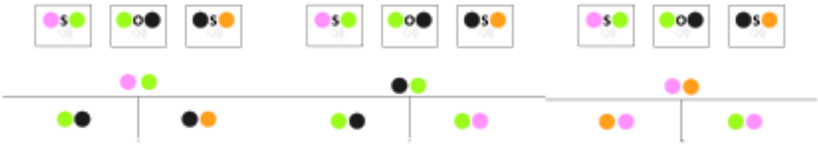
Example: Which one of these (points to comparisons) is like the one at the top (points to sample)?

Comparison



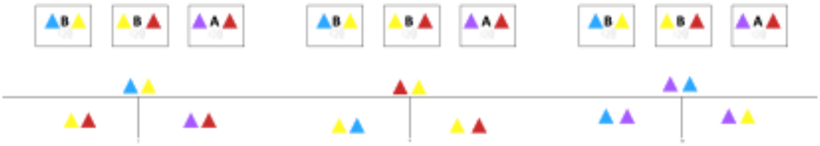
Example: Which one of these (points to comparisons) is like the one at the top (points to sample)?

Opposition



Example: Which one of these (points to comparisons) is like the one at the top (points to sample)

Temporality



Example: Which one of these (points to comparisons) is like the one at the top (points to sample)?

Hierarchy

Example: Which one of these (points to comparisons) is like the one at the top (points to sample)?

Appendix B

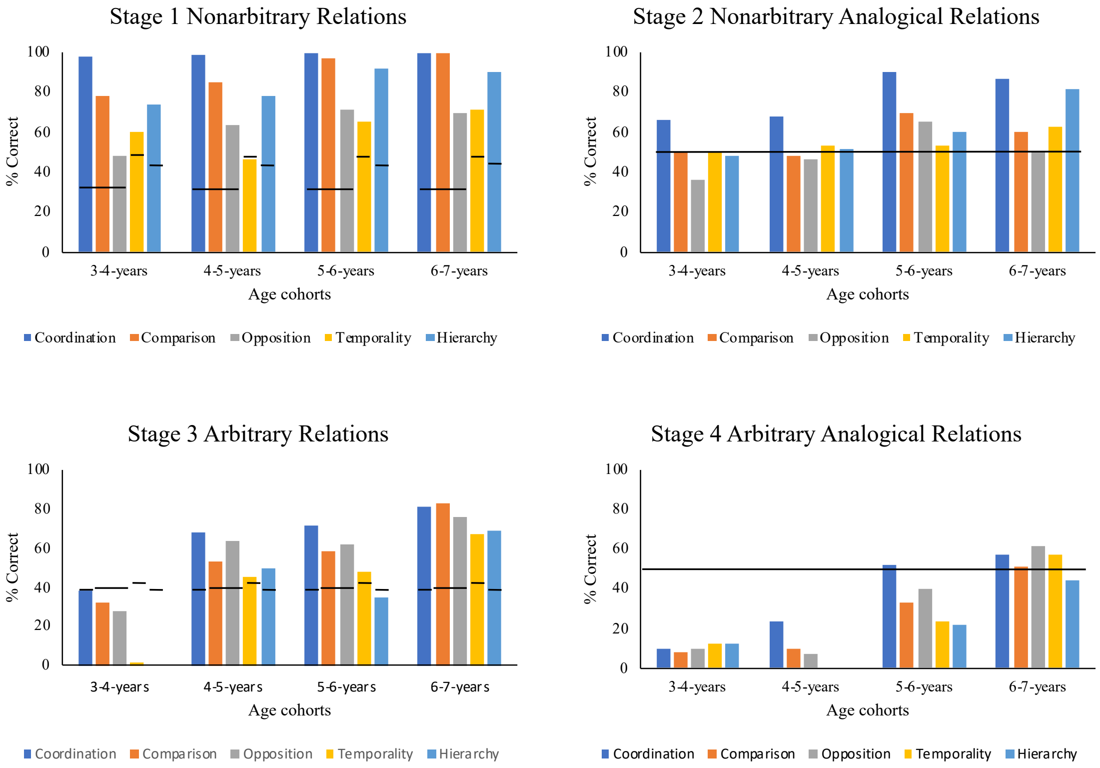
Table 1

Chance Level Responding Across Stages and Substages

	Stage 1	Stage 2	Stage 3	Stage 4
Coord	33%	50%	38%	50%
Comp	33%	50%	40%	50%
Opp	33%	50%	40%	50%
Temp	50%	50%	42%	50%
Hier	45%	50%	40%	50%

Note. This table represents chance-level responding, or what you might expect a child to get correct purely on the basis of chance or consistently guessing. These data can be used to compare with participant relational assessment scores.

Figure 1

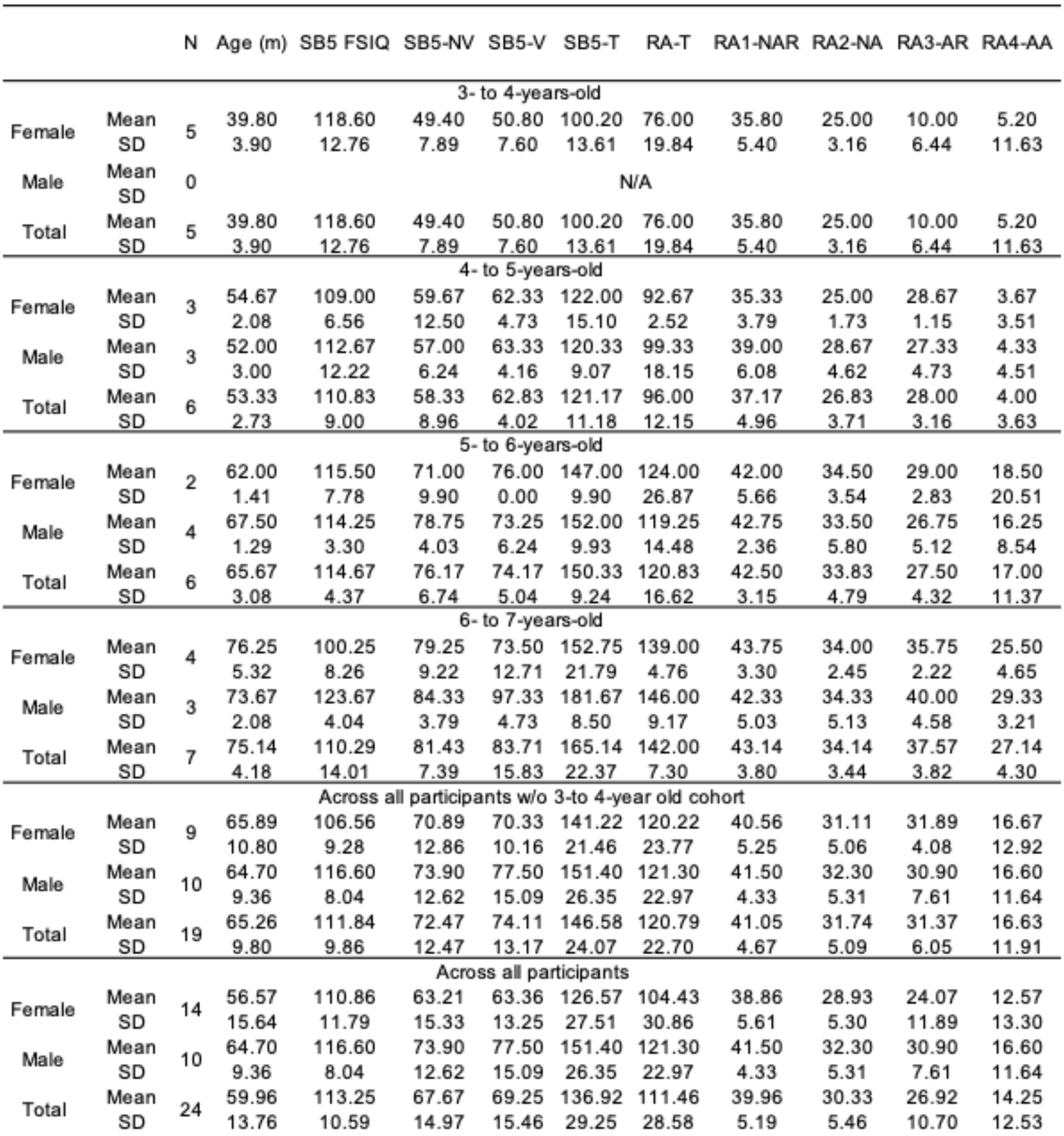


Note. Mean relational assessment scores across substages and age cohorts for each stage in the relational assessment.

Arguments against the chance level of responding: Children could be getting the simpler questions correct and the harder questions incorrect, which is suggested by participant improvement across the age groups. This table probably better represents harder relations where they really do not know the answer and are likely just guessing.

Appendix C

Analysis of Relational Assessment Performance by Gender

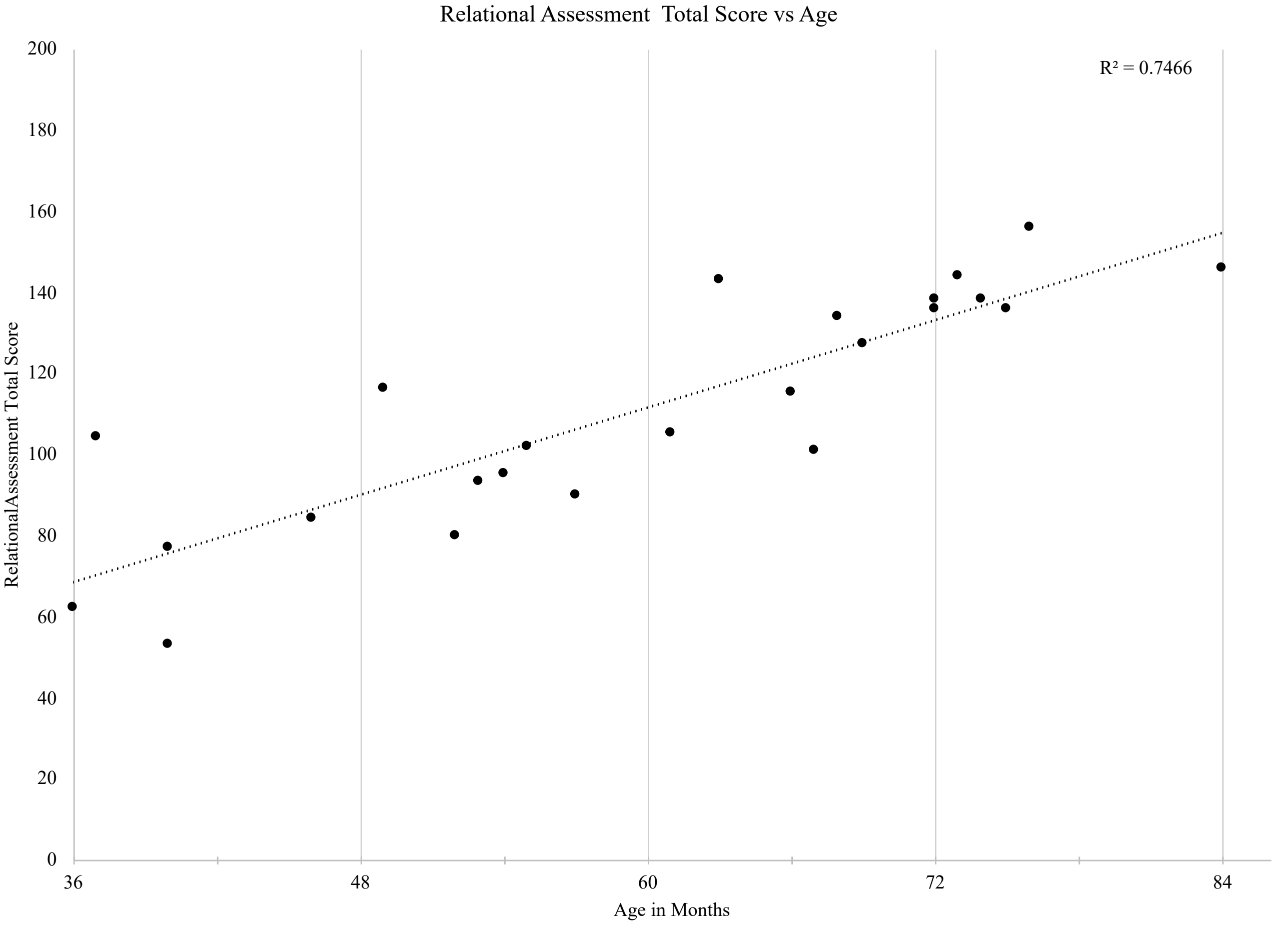


		N	Age (m)	SB5 FSIQ	SB5-NV	SB5-V	SB5-T	RA-T	RA1-NAR	RA2-NA	RA3-AR	RA4-AA
3- to 4-years-old												
Female	Mean	5	39.80	118.60	49.40	50.80	100.20	76.00	35.80	25.00	10.00	5.20
	SD		3.90	12.76	7.89	7.60	13.61	19.84	5.40	3.16	6.44	11.63
Male	Mean	0	N/A									
	SD											
Total	Mean	5	39.80	118.60	49.40	50.80	100.20	76.00	35.80	25.00	10.00	5.20
	SD		3.90	12.76	7.89	7.60	13.61	19.84	5.40	3.16	6.44	11.63
4- to 5-years-old												
Female	Mean	3	54.67	109.00	59.67	62.33	122.00	92.67	35.33	25.00	28.67	3.67
	SD		2.08	6.56	12.50	4.73	15.10	2.52	3.79	1.73	1.15	3.51
Male	Mean	3	52.00	112.67	57.00	63.33	120.33	99.33	39.00	28.67	27.33	4.33
	SD		3.00	12.22	6.24	4.16	9.07	18.15	6.08	4.62	4.73	4.51
Total	Mean	6	53.33	110.83	58.33	62.83	121.17	96.00	37.17	26.83	28.00	4.00
	SD		2.73	9.00	8.96	4.02	11.18	12.15	4.96	3.71	3.16	3.63
5- to 6-years-old												
Female	Mean	2	62.00	115.50	71.00	76.00	147.00	124.00	42.00	34.50	29.00	18.50
	SD		1.41	7.78	9.90	0.00	9.90	26.87	5.66	3.54	2.83	20.51
Male	Mean	4	67.50	114.25	78.75	73.25	152.00	119.25	42.75	33.50	26.75	16.25
	SD		1.29	3.30	4.03	6.24	9.93	14.48	2.36	5.80	5.12	8.54
Total	Mean	6	65.67	114.67	76.17	74.17	150.33	120.83	42.50	33.83	27.50	17.00
	SD		3.08	4.37	6.74	5.04	9.24	16.62	3.15	4.79	4.32	11.37
6- to 7-years-old												
Female	Mean	4	76.25	100.25	79.25	73.50	152.75	139.00	43.75	34.00	35.75	25.50
	SD		5.32	8.26	9.22	12.71	21.79	4.76	3.30	2.45	2.22	4.65
Male	Mean	3	73.67	123.67	84.33	97.33	181.67	146.00	42.33	34.33	40.00	29.33
	SD		2.08	4.04	3.79	4.73	8.50	9.17	5.03	5.13	4.58	3.21
Total	Mean	7	75.14	110.29	81.43	83.71	165.14	142.00	43.14	34.14	37.57	27.14
	SD		4.18	14.01	7.39	15.83	22.37	7.30	3.80	3.44	3.82	4.30
Across all participants w/o 3-to 4-year old cohort												
Female	Mean	9	65.89	106.56	70.89	70.33	141.22	120.22	40.56	31.11	31.89	16.67
	SD		10.80	9.28	12.86	10.16	21.46	23.77	5.25	5.06	4.08	12.92
Male	Mean	10	64.70	116.60	73.90	77.50	151.40	121.30	41.50	32.30	30.90	16.60
	SD		9.36	8.04	12.62	15.09	26.35	22.97	4.33	5.31	7.61	11.64
Total	Mean	19	65.26	111.84	72.47	74.11	146.58	120.79	41.05	31.74	31.37	16.63
	SD		9.80	9.86	12.47	13.17	24.07	22.70	4.67	5.09	6.05	11.91
Across all participants												
Female	Mean	14	56.57	110.86	63.21	63.36	126.57	104.43	38.86	28.93	24.07	12.57
	SD		15.64	11.79	15.33	13.25	27.51	30.86	5.61	5.30	11.89	13.30
Male	Mean	10	64.70	116.60	73.90	77.50	151.40	121.30	41.50	32.30	30.90	16.60
	SD		9.36	8.04	12.62	15.09	26.35	22.97	4.33	5.31	7.61	11.64
Total	Mean	24	59.96	113.25	67.67	69.25	136.92	111.46	39.96	30.33	26.92	14.25
	SD		13.76	10.59	14.97	15.46	29.25	28.58	5.19	5.46	10.70	12.53

Note. There are no substantive differences between males and females based on the small N, either within stages or in terms of the total relational score.

Appendix D

Relational Assessment Score vs. Age



Appendix E

Relational Assessment Score vs. SB5 Raw Score