Individual differences in the process of relational reasoning

Emily M. Grossnickle a, *, Denis Dumas b, Patricia A. Alexander b, Peter Baggett b

a Georgetown University, USA
b University of Maryland, College Park, USA

ARTICLE INFO

Article history:
Received 6 May 2015
Received in revised form 26 December 2015
Accepted 29 January 2016
Available online 11 February 2016

Keywords:
Relational reasoning
Bayesian network analysis
Componential model
Analogical reasoning
Working memory

ABSTRACT

The ability to discern meaningful patterns—relational reasoning—has been identified as a process important for student learning and cognition. Yet, research has typically investigated performance over processing, particularly when examining the role of factors such as working memory capacity. Moreover, studies have focused on analogical reasoning to the exclusion of other identified relational forms (i.e., anomaly, antinomy, antithesis). Study 1 investigates the role of individual differences in relational reasoning performance across four relational forms. Then, Study 2 identifies the highest and lowest-performing students from Study 1 to examine the probability that undergraduate students reach each of four sequential component processes of reasoning and the degree to which significant individual differences from Study 1 (i.e., visuospatial working memory) play a role in each process. Results indicate that low performing students experience particular difficulties in identifying relevant inferences and in mapping those inferences. This was due in part to the relation between working memory capacity and the processes of inferring and mapping. The outcomes from this study contribute to understandings of the sources of success and failure in reasoning for students at different levels, and identify potential entry points for intervention research.

1. Introduction

More than at any previous point in human history, today’s students inhabit an educational landscape awash in information (Lenhart, 2015; Purcell et al., 2012). As such, learners must work continually to make sense of divergent data, identify trends within seemingly unrelated ideas, and recognize patterns within and across domains (Alexander et al., 2011; Bråten & Strømsø, 2010). Although the processes and abilities necessary to support student success in such endeavors are expansive, this investigation focuses on one intriguing and powerful cognitive capacity, relational reasoning (e.g., Gentner, Loewenstein, & Thompson, 2003; Gick & Holyoak, 1980; Richland & McDonough, 2010).

Relational reasoning, the ability to discern meaningful patterns within any stream of information (Alexander & the DRLRL, 2012; Bassok, Dunbar, & Holyoak, 2012), has long been regarded as a foundational cognitive ability (James, 1890/1950; Sternberg, 1977), and is widely considered critical for learning and academic performance (e.g., Farrington-Flint, Canobi, Wood, & Faulkner, 2007; Schiff & Ravid, 2007). Although reasoning about patterns of similarity (i.e., analogical reasoning) have dominated the educational literature (Dumas, Alexander, & Grossnickle, 2013), additional relational forms have been identified. Specifically, reasoning about patterns of aberrance (i.e., anomalous reasoning), incompatibility (i.e., antinomous reasoning), and opposition (i.e., antithetical reasoning) have emerged as forms of relational reasoning important for learning in many domains (Broughton, Sinatra, & Reynolds, 2010; Chin & Malhotra, 2002; Kuhn & Udell, 2007; Sidney, Hattikudur, & Alibali, 2015).

Across various domains and levels of learning (e.g., Dumas, Alexander, Baker, Jalbinsky, & Dunbar, 2014; Ehri, Satlow, & Gaskins, 2009; Richland & McDonough, 2010), analogical reasoning has been found to correlate with a number of individual differences (Fleischhauer et al., 2010; Krawczyk, 2012; Mackintosh & Bennett, 2005; Morrison, Holyoak, & Truong, 2001), with working memory chief among them. Yet, studies examining individual differences in non-analogical reasoning have been limited. Moreover, relational reasoning has typically been examined as a capacity or ability measured through the outcomes of successful or unsuccessful performance on reasoning tasks. However, within the analogical reasoning literature, models acknowledge that successful
relational reasoning requires the execution of a sequence of component processes (e.g., inferring or mapping) rather than any singular or gestalt process (Krawczyk, McClelland, Donovan, Tillman, & Maguire, 2010; Sternberg, 1977). What is unknown, is the degree to which students successfully engage in the component processes ascribed to analogical reasoning when reasoning about other relational forms (e.g., analogy). Further, componental models propose that relational reasoning processes are sequentially dependent, that is, each preceding process must be completed in order to move on to the subsequent process (Lifshitz, Weiss, Tzuriel, & Tzemach, 2011; Sternberg, 1977). However, to our knowledge, no existing research examining the component processes of relational reasoning has utilized statistical methods examining the probabilistic dependence of these processes. The present study moves forward the statistical examination of the componental model of analogical reasoning by applying Bayesian network analysis to test the probability of successful completion of a given component process (e.g., mapping) predicated on the successful completion of the preceding component process (e.g., inferring).

These observed gaps in the literature have particularly important implications for educational practice and for the development of component-based training studies. For example, if a student appears unable to successfully reason relationally, the research literature currently offers limited empirically derived clues as to where their reasoning process may be breaking down or the degree to which other factors, such as working memory capacity, may be influencing specific components of the reasoning processes. Although previous studies have examined the effectiveness of component-based training for analogical reasoning (Alexander, White, Haensly, & Crimmins-Jeanes, 1987; White & Alexander, 1986), the extent to which these component processes apply to non-analogical forms of relational reasoning remains underexamined. Further, by investigating the degree to which individual differences such as working memory predict the component processes of analogical and non-analogical forms of relational reasoning, the present investigation provides insights into how these individual factors might support or hinder overall performance (Study 1) or the execution of specific components (Study 2). As such, inquiry into these gaps could prove diagnostically invaluable and become the basis for subsequent intervention.

Therefore, the present two studies focus on the identification and explication of specific junctures in the process of relational reasoning where individuals’ performance may falter. To do this, Study 1 assessed students’ performance on multiple forms of relational reasoning and related performance to relevant individual differences. Then, Study 2 identified high- and low-performing students from Study 1 and utilized conditional-probabilistic, or Bayesian, networks to uncover specific points in the reasoning process where certain students progress while others do not. Additionally, the significant individual differences from Study 1 were retained in Study 2 to examine how these characteristics related to critical points within the reasoning process.

2. Theoretical framework

2.1. Forms of relational reasoning

At its core, relational reasoning consists of identifying relations among relations, referred to as higher-order relations (Crone et al., 2009; Gentner, 1983; Krawczyk, 2012). Higher-order relations involve the identification of a pattern across seemingly disparate information (Cick & Holyoak, 1980; Goswami & Brown, 1990). This pattern is not simply the drawing of an inference between two ideas or objects (i.e., lower-order relations). Instead, it involves aggregating or mapping multiple lower-order relations in meaningful ways (Chi & VanLehn, 2012; Gentner, 1983; Holyoak & Thagard, 1989). Relational forms (i.e., analogy, anomaly, antinomy, antithesis) are characterized according to the type of higher-order patterns that are required (i.e., similarity, aberrance, incompatibility, opposition; Alexander & the DRLRL, 2012; Dumas et al., 2013). Although these higher order patterns are composed of lower-order patterns, it is the higher order patterns that distinguish the different forms of relational reasoning.

Clearly, the most commonly studied higher-order relation in the research literature is analogy, which requires a higher-order relation of similarity (Alexander, Dumas, Grossnickle, List, & Firetto, 2015). For example, a cell may be said to be relationally similar, or analogically related to, a factory because both are parts of a larger collective (i.e., a body or society) and contain interdependent parts with similar functions (e.g., the nucleus serves as the headquarters or manager). However, higher-order relations other than similarity have emerged within the theoretical and empirical literature; anomaly, antinomy, and antithesis (Dumas et al., 2013; Ferguson & Sanford, 2008; Stewart, Kidd, & Haigh, 2009; Tanca, Grossberg, & Pinna, 2010). Although these are not argued to be the only forms, they have been put forward as forms important for learning and development (Alexander et al., 2015; Dumas et al., 2012).

In contrast to analogy, which requires a higher-order relation of similarity, anomaly involves the identification of cases that are aberrant and, thus, do not fit within an overarching scheme (Chinn & Brewer, 1993; Klahr & Dunbar, 1988; Kulkarni & Simon, 1988). Reasoning by anomaly requires the identification of a higher-order relation of discrepancy or deviation between the anomalous idea, object, or event and the others. To do this various lower-order patterns are identified among multiple objects or concepts within a body of information, and then a higher-order relation of discrepancy is mapped. For example, when statisticians work to identify outliers in a regression analyses, they must first consider the relations among all their data points in order to conceptualize a particular pattern (e.g., the regression line), in which most but not all of the cases fall.

A third form of relational reasoning, antinomy, requires the conceptualization of mutual exclusion between or among ideas (Chi & Roscoe, 2002; Cole & Wertsch, 1996; Sorensen, 2003). In this way, antinomy requires the mapping of a higher-order relation of incompatibility among ideas, objects, or events whose lower-order relations have been characterized. For example, when a medical student learns to make accurate diagnoses, they practice “ruling-out” possible conditions by attending to the symptoms a patient is presenting and deciding whether those symptoms are compatible or incompatible with a given diagnosis (Dumas et al., 2014).

Finally, reasoning via antithesis involves the identification of a higher-order pattern of opposition between concepts (Bianchi, Savardi, & Kubovy, 2011; Kuhn & Udell, 2007). For example, in the educational context students may encounter refutation texts that present two directly opposing viewpoints on a given topic, such as climate change, which require students to conceptualize higher-order relations of opposition among the various lower-order relations in each argument (Sinatra & Broughton, 2011). Study 1 investigates the degree to which relevant individual differences (e.g., working memory capacity, need for cognition) are associated with performance on each of the forms of relational reasoning. Then, Study 2 examines the componental process of each of the forms of relational reasoning, with an eye towards identifying the point where that process collapses for low-performing students, and explaining that collapse with individual difference variables from Study 1. Additionally, Study 2 compares the degree to which students are more or less successful in executing the processes of relational reasoning across each of the four forms.
2.2. Processes of relational reasoning

Precisely because relational reasoning is held to consist of a number of discrete, sequentially related processes contributing to successful performance, this cognitive undertaking can be characterized as componential. In fact, there has been theoretical and empirical work into analogical reasoning that has supported this contention (Carpenter, Just, & Shell, 1990; Hummel & Holyoak, 2005; Lifshitz et al., 2011; Sternberg, 1977). What we sought to do in Study 2 was to build on this process-oriented approach in analogy research, particularly Sternberg’s componential model of analogical reasoning, to explore the componential nature of each of the four forms of relational reasoning that frame this study. Given the success of training component processes in analogical reasoning (Alexander et al., 1987; White & Alexander, 1986), examining how these processes unfold for other forms of relational reasoning is of particular importance.

2.2.1. Component processes

As noted, analogical reasoning has been conceptualized as requiring a series of component processes (i.e., encode, infer, map, and apply) that individuals sequentially complete in order to successfully construct a higher-order relation among pieces of information (Sternberg, 1977; see Table 1). For instance, the relevant features of the problem must first be identified or encoded (Carpenter et al., 1990; Hummel & Holyoak, 2005). In effect, the first step in reasoning analogically is to recognize and comprehend the individual elements that are parts of the problem space, that is, to encode the individual objects and their constituent parts. Take the classic analogy problem, [hand:fingers::foot:toe], as a case in point. The individual attempting to solve this particular problem must attend to the individual terms (i.e., hand, fingers, and foot) and comprehend the meaning of each (i.e., encode).

Table 1
Examples of processes of relational reasoning across relational forms.

<table>
<thead>
<tr>
<th>Reasoning task</th>
<th>Encode</th>
<th>Infer</th>
<th>Map</th>
<th>Apply</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analogy</td>
<td>Identify the relevant features of a problem</td>
<td>Identify lower-order relations between the individual features</td>
<td>Identify higher-order pattern between the lower-order relations</td>
<td>Determine a correct response to the task</td>
</tr>
<tr>
<td>Analogy</td>
<td>Identify the parts of a cell (e.g., nucleus, cell membrane) and the parts of a factory (e.g., headquarters, assembly line)</td>
<td>Identify the nucleus directs the actions of the cell; Identify that the headquarters directs the actions of a factory</td>
<td>Identify that the nucleus serves the cell in the same way that the headquarters serves a factory</td>
<td>A 7th grade biology student writes a paragraph accurately describing the relation between a cell and a factory</td>
</tr>
<tr>
<td>Anomaly</td>
<td>Review individual participants’ data for the predictor and criterion variables</td>
<td>Consider relations between an individuals’ data point and the average Calculate the mean based on set of data points; Examine a plot of the data against a regression line</td>
<td>Identify an aberrance in the pattern between the inferences Identify that most data points stay within 2SD of the mean, but three fall outside that range</td>
<td>A graduate student identifies outliers in her data to be removed</td>
</tr>
<tr>
<td>Antinomy</td>
<td>Identify the relevant parts of a medical case List symptoms patient is presenting; List symptoms of the diagnosis in question</td>
<td>Consider the relation between a patient’s symptoms and the diagnosis presented Identify that Disease A explains three of the patient’s main symptoms</td>
<td>Identify an incompatibility between the symptoms of the patient and the presented diagnosis Recognize that although the patient has some of the symptoms of Disease A, he does not have a fever or dizziness, therefore his symptoms are not compatible with Disease A</td>
<td>A medical student explains to her team that a patient’s symptoms are incompatible with the given diagnosis</td>
</tr>
<tr>
<td>Antithesis</td>
<td>Read the claims presented in two articles about climate change Read claims in Article A supporting the role of humans in climate change; Read claims in Article B denying the effect of human impact on climate change</td>
<td>For each article separately, consider how the claims relate Recognize that the author of Article A presents evidence that rising sea water was caused by human activity; Recognize that the author of Article B is presents evidence that rising sea water was not caused by human activity</td>
<td>Recognize that the findings related to sea level rise described in the two articles are depicted as being caused by opposite processes</td>
<td></td>
</tr>
<tr>
<td>Antithesis</td>
<td>A student in a high school environmental science class correctly identifies counterarguments in a class debate about climate change</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Once these terms are identified and characterized individually, associations between them can be inferred (Gentner, 1983, 2010). Comparable to the general reading literature, an inference is a lower-order association that entails the building of a relation within information by generating an understanding that is never expressly stated. If someone reads that Sam and Bill are stepbrothers, that reader could infer that Sam and Bill are not biologically related. For example, in the aforementioned analogy problem, conceptualizing the lower-order relation between hands and fingers must be derived (e.g., the hand contains fingers). Then, constituent lower-order relations are mapped into overarching patterns, often termed higher-order relations (Gick & Holyoak, 1980; Green, Fugelsang, & Dunbar, 2006). Returning to our classic analogy problem, the solver must formulate a relation between hand and foot (e.g., appendages that consist of digits) that constitutes a higher-order relational representation. After mapping is complete, the discerned higher-order pattern can be applied to determine a correct response to given task or problem (Sternberg, 1977). For the analogy problem, application would require identifying a component of the foot that parallels as close as possible the relation that fingers have to the hand (e.g., digits that are critical elements of the appendage and promote its functioning: toes). Although the aforementioned terms are in keeping with Sternberg’s (1977) componential model, similar but differentially-labeled processes have been identified by others delving into analogical reasoning (Carpenter et al., 1990; Gentner, 1983; Krawczyk et al., 2010). For instance, Carpenter et al. (1990) described the analogical reasoning process as incremental problem solving, and referred to perceptual analysis (encoding) and generalizing (mapping).

Sternberg’s (1977) model was originally designed to explain the process of analogical reasoning. However, given that relational reasoning in all forms constitutes higher-order relations, it is likely that the component processes in analogical reasoning are similar to...
the processes required to reason anomalously, anomalously, and antithetically. For instance, regardless of the presentation method (e.g., verbal or graphical), the relevant information must be encoded, lower-order relations inferred, higher-order relations mapped, and finally applied in order to correctly solve any relational reasoning problem (Chi & VanLehn, 2012; Gick & Holyoak, 1980). Indeed, since Sternberg's conceptualization, these processes have been utilized to explain reasoning and transfer (e.g., Chi & VanLehn, 2012; Gentner, 2010; Gentner & Rattermann, 1993; Lifshitz et al., 2011). However, studies explicitly addressing these processes during other forms of relational reasoning have been limited, and the processes have never, to our knowledge, been empirically examined for multiple forms of relational reasoning in the context of the same study. The present study expands on this prior research by comparing how these processes unfold for students engaging in multiple forms of relational reasoning, namely, analogy, anomaly, antinomy, and antithesis.

To illustrate how these component processes might unfold across each of the relational forms, Table 1 examines the four educational examples presented to describe the relational forms in Section 2.1. As is evident in the summary of examples across each of the relational forms, encoding involves the identification of relevant aspects of the problem, and inferring requires connections to be made among the encoded features. Then, it is the stage of mapping where the forms most notably diverge, as lower-order inferences are combined to create the particular higher-order pattern associated with the relational form (i.e., similarity, aberrance, incompatibility, opposition). The application of reasoning with these varied relational forms can take many instantiations across academic domains; however, we have provided one example for each form to illustrate how the mapped relations might be applied.

These component processes are logically determined to follow a particular sequence (i.e., encode, infer, map, apply) with the preceding process being necessary for individuals to complete the subsequent processes (Lifshitz et al., 2011; Sternberg, 1977). For example, without encoding at least some of the relevant information, it is impossible for a participant to form an inference. In the same way, because mapping is defined as the formation of a higher-order relation among multiple inferences, the formation of at least some of those inferences must logically precede mapping. Finally, because applying in this case is defined as the utilization of a mapped relation to generate or select an answer (and not simply guess) mapping must logically precede the application of the mapped relation. Of course, some recursion is possible, such as going back and encoding more of a problem after making a first inference (Carpenter et al., 1990; Sternberg, 1977), but logically, at least one instance of an earlier phase must precede the beginning of a later one. Even if these phases unfold too rapidly for a participant to be conscious of them, higher-order patterns cannot be mapped if item features have not been encoded, and relevant lower-order relations have not been inferred (Carpenter et al., 1990; Lifshitz et al., 2011).

2.2.2. Use of Bayesian network models to study component processes

Following the assumptions of the componental model of relational reasoning (i.e., that the processes unfold in a sequence with earlier phases preceding later ones), Study 2 utilizes Bayesian network analysis to examine the systematic unfolding of the componental processes. In contrast to statistical methods used to compare groups or estimate model fit, Bayesian network analysis estimates probabilistic networks among events (Neapolitan, 2004). In its simplest form, Bayesian network analysis can be used to estimate conditional probabilities; that is, the likelihood that event B will occur given that event A has already occurred, or P(B|A). This methodology is particularly well suited for research endeavors where the likelihood of a given outcome occurring is dependent on the outcome of temporally previous events. In education, Bayesian network analyses have received increasing attention given their potential for estimating the likelihood of educational events that are believed to be sequential, developmental, or contingent upon prerequisite skills (Schultz, 2007). This is of particular importance when considering learning that may unfold across development or examining the point at which individuals are likely (or unlikely) to progress in their learning. For example, Bayesian network analyses have been used for a variety of educational problems, such as estimating the likelihood of a student being able to generalize the meaning of a novel word (e.g., Tennenbaum, Griffiths, & Kemp, 2006) and modeling the probability that students will succeed in an online course as dependent on their engagement with online course features (Garcia, Amandi, Schiaffino, & Campo, 2007).

In the present investigation (Study 2) Bayesian network analysis was chosen in lieu of other statistical techniques because it allows for the examination and comparison of conditional probabilities (e.g., the likelihood of mapping if an inference was made). Given that Sternberg’s (1977) componental model forwards that earlier phases must precede a later one, Bayesian network analysis was well suited to the present investigation, where differences in conditional probabilities across groups of individuals and relational forms was an empirical focus. For this work, Bayesian methods were preferable to more traditional methods of calculating the probability that individuals map higher-order relations independent of whether they made inferences (i.e., the unconditional probability of mapping), or comparing average performance on the TORR scales. Indeed, Bayesian methods have been fruitfully applied to research questions within the relational reasoning literature (e.g., Holyoak, Lee, & Lu, 2010), and therefore hold promise for elucidating the dependency among the component processes examined in the present investigation.

2.3. Individual differences and relational reasoning

Prior research has indicated that the sequence of incremental processing associated with analogical reasoning has been found to be consistent across individuals (Carpenter et al., 1990; Sternberg, 1977). Yet, the degree to which students are able to successfully enact the various component processes has been found to vary according to certain individual differences. For instance, in a study of individuals with intellectual disabilities, Lifshitz et al. (2011) found that compared to individuals with moderate disabilities, those with mild disabilities performed significantly better across almost all phases (except mapping) when completing conceptual and perceptual figural analogues. Moreover, in a comparison of average and better performers on a test of analogical reasoning, better performers were able to identify more complex correspondences between two figures and to induce more abstract rules (Carpenter et al., 1990). At the same time, if average performers were able to induce patterns and generalize rules (map), they were similar to the better performers in their ability to generate a solution.

Nevertheless, most of the research examining individual differences in relational reasoning has focused on analogical reasoning, specifically, analogical reasoning ability. The focus, therefore, has typically been the relation between various individual difference variables (e.g., working memory, gender) and performance on analogical reasoning tasks (e.g., Cho, Holyoak, & Cannon, 2007; Richland & Burchinal, 2013). However, like analogical reasoning, other relational forms (e.g., anomaly, antinomy,
antithesis) are characterized by the identification of higher-order relations (Alexander & the DRLRL, 2012; Dumas et al., 2013). If certain individual difference variables (e.g., working memory) are important for identifying higher-order relations, then the relations between specific individual differences and analogical reasoning should replicate for non-analogical forms of reasoning. Yet, if there were something particular to the identification of higher-order relations of similarity compared to other forms of higher-order relations, then it would be expected that individual differences would predict some forms of relational reasoning but not others. Therefore, Study 1 examines whether certain relevant individual differences (i.e., working memory capacity, need for cognition, gender) significantly relate to multiple forms of analogical and non-analogical relational reasoning performance. Then, to examine how relevant individual differences impact performance, Study 2 investigates the influence of individual differences on each of the four component processes. In the present investigation, working memory capacity, need for cognition, and gender were selected to represent cognitive and non-cognitive characteristics commonly associated with visuospatial analogical reasoning, which have the potential to influence relational reasoning in its multiple forms.

2.3.1. Working memory capacity (WMC)

Important for numerous academic tasks, working memory involves the “temporary storage and manipulation of information” (Baddeley, 1992, p. 556). Described as a coordinated system, working memory includes an executive control in addition to separate subsystems for the maintenance and manipulation of verbal/phonological and visuospatial information (Baddeley, 1992, 2001). Phonological working memory includes the storage and manipulation of auditory information, as well as written information that can be maintained in memory by subvocal repetition (Baddeley, 1992). This can include visually presented words, non-words, letters, or numbers (Baddeley, 2001; Gathercole & Adams, 1993). Visuospatial working memory, on the other hand, involves the storage and manipulation of visual patterns and spatial locations (Baddeley, 1992; Logie, 1995).

Given these separate subsystems, phonological and visuospatial working memory are typically measured using separate tasks to assess the maintenance of verbal and visuospatial information (e.g., Alloway, Gathercole, Kirkwook, & Elliott, 2008; Sprenger et al., 2013). Moreover, practice within the field of cognitive science suggests that studies may benefit by measuring each subsystem of working memory capacity with multiple tasks (Foster et al., 2015).

The focus of the present studies was to understand how individual differences in visuospatial working memory capacity related to the performance and process of relational reasoning. Given the potential importance of visuospatial WMC in relation to visuospatial relational reasoning as measured in the present study, we utilized two tasks to measure the visuospatial subsystem of WMC (i.e., Block Span, ShapeBuilder). As phonological working memory has not been found to significantly impact performance on visual-spatial analogical reasoning tasks (Rao & Baddeley, 2013), a measure of phonological WMC was included in the present study as a measure of discriminant validity.

A robust line of research has suggested that WMC is deeply related to analogical reasoning performance, even when controlling for factors such as age and processing speed (e.g., Cho et al., 2007; Fry & Hale, 1996; Krawczyk et al., 2008; Richland & Burchinal, 2013; Waltz, Lau, Gruwal, & Holyoak, 2000). Indeed, WMC and the ability to process information effectively in working memory may serve as the source for many of the differences in relational reasoning performance (Carpenter et al., 1990; Morrison et al., 2001; Rao & Baddeley, 2013). For instance, working memory has been implicated as a mechanism by which students keep track of specific aspects of a given problem, as well as their problem-solving goals (Carpenter et al., 1990)—a challenging, yet critical aspect of relational reasoning.

Given the need to keep track of many parts of a problem, the ability to process information in working memory may also be particularly critical as relational complexity (i.e., the number of relations) and problem difficulty increase (Birney & Halford, 2002; Cho et al., 2007). As the number of lower-order relations within a problem increases, the load on WMC also increases, because each inference must be held in working memory in order to be integrated. For instance, previous studies found that brain regions associated with visuospatial working memory and executive control of working memory were more active when individuals solved difficult compared to easier analogy items (Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997).

The present set of studies examined the relation between WMC and relational reasoning by investigating: (a) the degree to which higher WMC is related to successful performance of relational reasoning problems across multiple relational forms (Study 1), and (b) what point in the componential process of reasoning WMC is the most important (Study 2).

2.3.2. Need for cognition (NFC)

In relation to analogical reasoning, NFC has been found to serve as a contributor unique from working memory (Hill et al., 2013); however, its role has been less consistent, and deserves additional inquiry (Day, Espejo, Kowollik, Boatman, & McEntire, 2007; Fleischhauer et al., 2010). Need for cognition, often described as both a motivation and a personality trait, captures the degree to which individuals seek out and enjoy opportunities to engage in effortful cognitive endeavors (Caccioppo & Petty, 1982; Caccioppo, Petty, & Kao, 1984). As a motivational variable, need for cognition is regarded as domain-general and intrinsic (Caccioppo & Petty, 1982). Individuals high in NFC tend to report more positive experiences and a willingness to spend more effort and time engaging in cognitively demanding tasks (Caccioppo, Petty, Feinstein, & Jarvis, 1996; Enge, Fleischauer, Brocke, & Stroebele, 2008; Verplanken, Hazenberg, & Palenewen, 1992). Research investigating the role of NFC in relational reasoning has focused on analogical reasoning, and has identified significant, albeit modest, positive relations (Day et al., 2007; Fleischhauer et al., 2010; Hill et al., 2013).

The characteristics of NFC make it well suited for exploration as a factor contributing to multiple forms of relational reasoning. NFC was specifically chosen from possible motivational variables because theoretical and empirical evidence related to NFC aligned well with the nature of the relational reasoning task used in the present study. Specifically, given the challenging nature of the task, we hypothesized that the desire to engage with complex ideas would provide students with the motivation to persist despite the challenge of the task. When reasoning about problems higher in relational complexity, multiple inferences must be made in order to map a higher-order pattern (Birney & Halford, 2002). The process of encoding and probing the available information for multiple inferences requires a willingness to effortfully engage with the problem at hand; a characteristic of individuals high in NFC (Caccioppo et al., 1996; Enge et al., 2008). Moreover, relational reasoning requires that individuals are willing to engage in the revision of inferences if their original patterns do not hold (Sternberg, 1977). Given the evidence that individuals higher in NFC are more willing to seek out additional information (Curseu, 2011) and consider opposing positions or shift beliefs (Kardash & Scholes, 1996; Murphy, Holleran, Long, & Zeruth, 2005), there is a strong theoretical rationale for empirically examining this relation.
2.3.3. Gender

Within the visuospatial cognition literature, males have been found to consistently outperform females in tests of visuospatial reasoning (e.g., Cooke-Simpson & Voyer, 2007; Debelek, Gittler, & Arendasy, 2014). These effects are particularly robust when students are required to make spatial manipulations such as mental rotation (Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995). These differences are of particular interest in the present study, given the visuospatial nature of the relational reasoning task, which includes items requiring the mental manipulation of visuospatial information. Given that visuospatial reasoning has significant implications for student performance in academic domains that emphasize visuospatial skills, such as the sciences, mathematics, and engineering (Uttal & Cohen, 2002), further investigation of gender differences is warranted.

Within the relational reasoning literature, gender differences in visuospatial relational reasoning performance have been identified, with males outperforming females (Mackintosh & Bennett, 2005), particularly in late adolescence and adulthood (Lynn & Irving, 2004). However, other research has indicated that gender differences in relational reasoning are not as prevalent as earlier research had suggested (Ruggiero, Sergi, & Iachini, 2008). Explanations for differences in visuospatial reasoning across males and females have implicated a number of different task-specific and individual factors. For instance, in a study of visuospatial analogical reasoning, Mackintosh and Bennett (2005) found superior performance in males across items ranging in difficulty; however, the most pronounced effects were evident for the most difficult items. Other individual factors such as working memory have the potential to explain gender differences given that males tend to outperform females on measures of visuospatial working memory, particularly under conditions of high demand (Kauffman, 2007; Ruggiero et al., 2008). The present study investigates whether gender is a unique contributor to visuospatial relational reasoning performance when working memory is taken into account.

3. The present studies

Two studies were conducted to examine the role of individual differences in relational reasoning. Study 1 examined whether individual differences identified as important for analogical reasoning are also related to other forms of relational reasoning performance. Specifically, Study 1 addressed the following research question (RQ1.1): To what degree are verbal and visuospatial working memory capacity, need for cognition, and gender related to relational reasoning performance for multiple relational forms?

Based on the findings of Study 1, Study 2 employed a selective sampling strategy to identify the highest and lowest performing students from Study 1 to examine how the component processes of relational reasoning unfold across multiple relational forms. These processes were examined for high- and low-performing students, and their success in each of the processes was examined in relation to the individual characteristics identified as important in Study 1. A think-aloud methodology was applied during students’ reasoning to examine the four component processes (i.e., encode, infer, map, apply) as they emerge in multiple forms of relational reasoning (i.e., analogy, antimony, antithesis). In Study 2, Bayesian network analyses were used to analyze the conditional likelihood that students were able to transition between each of the component processing phases, given that they had successfully engaged in the preceding phase. These conditional likelihoods were compared across levels of performance, problem difficulty, and relational reasoning form. Further, any significant individual differences from Study 1 were then examined in relation to these probabilities in order to investigate the degree to which individual differences matter across the phases of processing.

Specifically, five research questions were addressed in Study 2. First, RQ2.1 asked the following: Is the transition from one component process to the next less likely for certain processes? This question combined results for all Study 2 participants across all forms of relational reasoning to address whether there is a difference in the likelihood that students reached each processing phase (e.g., infer, map) given that they have engaged in the proceeding processes. Second, RQ2.2 was: Do high and low performers differ in the conditional likelihood that they reach given processing phases? This analysis separated and compared the results of RQ2.1 by high and low performers. The third research question (RQ2.3) was: Does the conditional likelihood that individuals reach a given processing phase differ by overall item difficulty? Problem difficulty for each item was characterized a priori based on the results from the larger pool of students in Study 1. This allowed us to investigate whether differences in problem difficulty might be due in part to challenges from a specific processing phase. For instance, do students struggle with difficult problems because they are unable to map or apply, or does the challenge occur earlier, such as during the inference phase? Fourth, RQ2.4 examined whether the conditional likelihood that individuals reach a given processing phase differs by form of relational reasoning. This question addressed whether analogy, anomaly, antimony, and antithesis items were similar or different for each of the processing phases (i.e., infer, map, apply). Finally, the fifth research question (RQ2.5) was: To what degree are individual differences identified as significant in Study 1 (i.e., visuospatial WMC) related to the conditional likelihood that individuals reach given phases of processing? The results and conclusions of the two studies are presented individually, followed by a general discussion of the conclusions and implications across the two studies.

4. Study 1

4.1. Study 1 method

4.1.1. Study 1 participants

Participants were 71 undergraduate students at a large mid-Atlantic American university (46 female; 63.9%) recruited from a human development course. In exchange for their participation, they were offered extra credit in that course. Participants ranged in age from 19 to 26, with a mean age of 21.10 (SD = 1.09). The sample was majority White (n = 57; 80.30%), with 4.2% Hispanic/Latino (n = 3); 14.3% African American (n = 1); 9.9% Asian or Pacific Islander (n = 7); and 4.2% participants reporting their ethnicity as other (n = 3). The vast majority (95.71%) of the sample reported their first language as English (n = 67), and participants reported a mean grade point average (GPA) of 3.39 (SD = .38) with GPAs ranging from 2.20 to 4.00.

4.1.2. Study 1 measures

This study included five measures: (a) the Test of Relational Reasoning (TORR), (b) the Need for Cognition Scale-Short Form (NCS), (c) Block Span, (d) Shapebuilder, and (e) Letter-Number-Sequencing (LNS). Gender was self-reported by participants. All of the measures were completed online during a regularly scheduled class meeting, under the supervision of one of the researchers.

4.1.2.1. Test of Relational Reasoning. In order to assess participants’ ability to discern meaningful patterns across multiple forms of relations, we administered the TORR (Alexander et al., 2015). The TORR is a 32-item relational reasoning test, designed to limit the influence of prior knowledge and language ability through the use of graphical, non-linguistic items. It has demonstrated appropriate reliability and validity with undergraduate samples (Alexander & Petersen, 1985; Voyer, Voyer, & Bryden, 1995).
et al., 2015), and internal consistency in the present study was high ($\alpha = .83$). This measure comprises four scales representing each of the four forms of relational reasoning previously described (i.e., analogy, analogy, analogy, and antithesis; see Appendix A for sample items). Each scale consisted of two practice items followed by eight test items, each worth 1 point. Total scores were calculated separately for each scale and for the total measure, with a possible range of 0–8 points per scale, and 0–32 points for the total measure. Supporting the construct validity of the TORR in the present study, and consistent with previous psychometric work (Alexander et al., 2015), moderate correlations were observed among the scales of the TORR ($r = .25$–.47), suggesting that the four forms of reasoning were positively related.

4.1.2.2. Need for Cognition Scale. The 18-item NCS short form ($\alpha = .89$ in the present study) was developed by Cacioppo et al. (1984) and asks participants to indicate their level of agreement or disagreement with statements such as “I find satisfaction in deliberating hard and for long hours” and “The notion of thinking abstractly is appealing to me” (Cacioppo et al., 1984, p. 307). Responses are recorded using a 9-point Likert scale from $-4$ (very strong disagreement) to $+4$ (very strong agreement), and half the items are reverse scored.

4.1.2.3. Block span. A computerized version of the widely used block span task (Kessels, van Zandvoort, Postma, Kappelle, & de Haan, 2000) was used as a measure of visuospatial working memory. Block span requires participants to remember the spatial location (in serial order) of a sequence of black blocks appearing for $1 \text{s}$ in a $4 \times 4$ grid. The Block span task score was computed by counting the number of blocks recalled in the correct serial order and spatial position.

4.1.2.4. Shapebuilder. Shapebuilder (Atkins et al., 2014; Sprenger et al., 2013) is a visuospatial working memory task in which participants are asked remember the order and spatial position that a series of colored shapes presented one at a time within a $4 \times 4$ grid. Participants observed a sequence of two, three, or four shapes (circles, triangles, squares, or diamonds) that were one of four colors (red, blue, yellow, or green), and were asked to recreate the sequence. The Shapebuilder task increases in difficulty by including more stimuli, and including a greater variety of colors and shapes. For example, at the easiest level, items were all the same shape or color, but appeared in different locations. At the most difficult level, items were all different colors and shapes, and appeared in different locations. Participants received 15 points for getting the first item of a trial correct, 30 points for the second, 60 for the third, and 120 for the fourth. If an item was incorrect, possible scoring started over on the next item. Points for partial accuracy were awarded such that five points were given for correct location only and 10 points were given for correct location and shape. The total points participants received for 26 trials represented the overall Shapebuilder score.

4.1.2.5. Letter number sequencing (LNS). The LNS (Alloway et al., 2008), a phonological working memory task, requires participants to remember a series of 1–4 letters (uppercase A through Z) and 1–4 numbers (numerals 1–9) randomly presented one at a time on a computer screen, for 500 milliseconds each. Each series consists of 2–12 serially presented letters and numbers. Following each sequence of letters and numbers, participants use the keyboard to enter the numbers in forward numerical order and the letters in forward alphabetical order, regardless of the order in which they were originally presented. For example, if a participant was presented with the series, “S D 9 C 2 J,” the correct response would be 2 5 9 in the number category and C D J in the letter category. Within a set, any given letter or number appeared only once. The LNS score was calculated using the number of correctly remembered and reordered characters across all presented series. Because LNS utilizes letters and numbers and engages manipulation through subvocal rehearsal, it is regarded as a phonological working memory measure, rather than a visuospatial one (Baddeley, 1992).

Given that the two measures of visuospatial WMC (i.e, Shapebuilder and Block Span) were highly correlated, a principal components analysis (PCA) was used to confirm that a single component could account for the variance in both measures. A single component, visuospatial working memory (VS-WMC), accounted for $87.77\%$ of the variance ($\lambda = 1.75$) in the Shapebuilder and Block Span tasks. PCA scores for VS-WMC were calculated for each participant in order to alleviate potential problems to later analysis caused by this observed collinearity, and to test the predictive power of visuospatial WMC generally. It should be noted that because the LNS was not a measure of VS-WMC, and as such correlated only moderately with Shapebuilder and Block Span, it was not included in the principal component and was analyzed separately.

4.1.3. Study 1 procedures

Data were collected during a regularly scheduled meeting of the students’ human development course. During this class meeting, the students used computers to independently complete each of the measures. After receiving a consent form, they were given the entire class period (75 min) to complete the measures and a demographics form. No student required more than the allotted 75 min to complete the measures.

4.2. Study 1 results

We addressed the research question for Study 1 by examining the degree to which gender and the individual differences of WMC and NFC were related with relational reasoning performance across four forms of relational reasoning (see Table 2). Pearson correlations were calculated for the relations between each of the variables except gender, which were calculated using a point-biserial correlation. The VS-WMC component was positively related to overall TORR performance as well as performance on each of the four TORR scales. However, LNS did not significantly correlate with the total TORR score or any of its scales. This finding is potentially explained by the non-spatial nature of the LNS and the highly spatial nature of the TORR. The different relations between the TORR and the various working memory measures provides support for convergent and discriminant validity of the TORR in the present study. It is notable that VS-WMC was also correlated with gender, indicating that males tended to receive higher scores on VS-WMC. In terms of relational reasoning, gender was positively related only to the analogy scale and to overall TORR performance, but not to the anomaly, antonymy, or antithesis scales. Finally, it is important to note that NFC did not correlate significantly with the TORR or any of its scales.

Further, to probe the degree to which individual differences were associated with relational reasoning ability when controlling for each of the other characteristics, a multiple regression analysis was undertaken. Specifically, VS-WMC, LNS, NFC, and gender were entered into a multiple linear regression model predicting TORR scores. The overall model was significant $F(4,66) = 9.12, p < .001$, $R^2 = .60$, $R^2 = .36$, and indicated that the individual differences explained more than a third of the variance in TORR performance. However, when considering the relation of each of the independent variables when all other variables in the model were held constant, only VS-WMC ability emerged as a significant predictor of
and phonological working memory were not significant. Moreover, it is potentially interesting to note that, while vsuospatial working memory was related to students’ performance on multiple forms of relational reasoning ability \[ t = 4.90, \beta = 2.93, \beta^* = .52, p < .001 \]. Therefore, VS-WMC was the only individual difference variable retained for analysis in Study 2.

4.3. Study 1 discussion

Study 1 examined the degree to which individual differences related to students’ performance on multiple forms of relational reasoning. Results indicated that visuospatial working memory was significantly related to all forms of relational reasoning performance, gender was related only to analogical reasoning, and NFC significantly related to students’ performance on multiple forms of relational reasoning. Results indicated that visuospatial working memory was related to students’ performance on multiple forms of relational reasoning. Moreover, it is potentially interesting to note that, while vsuospatial working memory had been established in the analogical reasoning literature, the present study revealed a similar relation to non-analogical forms of relational reasoning ability \[ t = 4.90, \beta = 2.93, \beta^* = .52, p < .001 \].

4.3. Study 1 discussion

Study 1 examined the degree to which individual differences related to students’ performance on multiple forms of relational reasoning. Results indicated that visuospatial working memory was significantly related to all forms of relational reasoning performance, gender was related only to analogical reasoning, and NFC and phonological working memory were not significantly related to any of the measured forms. Moreover, when controlling for visuospatial working memory, gender no longer accounted for a significant amount of variance in relational reasoning.

Although the importance of visuospatial working memory has been established in the analogical reasoning literature, the present study revealed a similar relation to non-analogical forms of relational reasoning, including anomalous, antinomous, and antithetical reasoning. Moreover, it is potentially interesting to note that, while VS-WMC was strongly associated with relational reasoning performance across all the relational forms, LNS did not significantly correlate with any of the scales of the TORR. Previous work (e.g., Baddeley, 1992) has found that the visuospatial and phonological components of working memory may be conceptually and empirically separable. Our findings support this argument by demonstrating that, given a visuospatial measure such as the TORR, VS-WMC was a significant predictor, while a phonological measure such as the LNS was not.

Finally, despite previous research indicating that NFC is positively related to analogical reasoning (Day et al., 2007; Fleischhauer et al., 2010; Hill et al., 2013), this relation was not significant in the present study. This held for all of the forms of relational reasoning. Although it is believed that the desire to engage in effortful processing indicative of NFC should increase performance on a range of tasks, the novel challenge of the TORR may have limited participants’ success at relational reasoning despite any desire to think in an effortful way. Moreover, the positive relation between NFC and analogical reasoning in some previous studies has been modest (e.g., Day et al., 2007; Fleischhauer et al., 2010; Hill et al., 2013), suggesting that the effect of NFC may not be as robust as measures such as visuospatial WMC.

5. Study 2

5.1. Study 2 method

5.1.1. Study 2 participants

Study 2 was completed using a subset of the participants from Study 1. Specifically, students who scored at least half of a standard deviation below or above the Study 1 mean on the TORR were recruited to participate in Study 2 at least two weeks after participating in Study 1. Students who scored \(0.5\ SD\) above the TORR mean in Study 1 are referred to as high performing (\(n = 16\)), and students who scored \(<0.5\ SD\) below the TORR mean in Study 1 are referred to as low performing (\(n = 16\)). In exchange for their participation in Study 2, students received additional extra credit in their human development course. Participants ranged in age from 19 to 26, with a mean age of 21.38 (SD = 2.126), and were 59.37% female (\(n = 19\)). The sample was majority White (\(n = 27\); 84.4%), with 31% Hispanic/Latino (\(n = 1\)); 9.4% Asian or Pacific Islander (\(n = 3\)); and 3.1% reporting their ethnicity as Other (\(n = 1\)). The majority (93.75%) of the sample reported their first language as English (\(n = 30\)), and participants reported a mean GPA of 3.45 (SD = .41) with GPs ranging from 2.40 to 4.00. Participants were classified as high or low performers based on their performance on the TORR in Study 1. Table 3 includes a breakdown of demographics and other individual difference variables by high and low performers.

To establish the difference between the high- and low-performing students, mean performance on the shortened 12-item version of the TORR given in Study 2 was compared (see Table 4). High-performing students answered 9.50 (SD = 1.41) items correctly on average, and low-performing students answered 5.56 (SD = 2.37) items correctly on average, \(t(30) = 5.72, p < .001\). The high-performing students also performed significantly better than the low-performing students on each of the four subscales (see Table 4 for averages and significance tests), suggesting that these empirically derived groups were stable across all four forms of relational reasoning. Moreover, the range of scores indicated that there was limited overlap between the overall performance of the high and low performers in Study 2. This suggests that performance was relatively stable from Study 1 to Study 2.
5.1.2. Study 2 measures
Participants completed a subset of the items of the TORR from Study 1, which included three items for each of the four scales. Items were selectively chosen based on their empirically observed item difficulty (p) levels in the full sample analyzed in Study 1. Specifically, one easy item (p > .70), one medium item (p = .31-.69) and one difficult item (p < .30) per scale were used. Item difficulty ranges such as these have been commonly utilized for the definition and identification of easy, medium, and difficult items on a variety of measures (Berk, 1986). In this way, three items from each of the four subscales of the TORR (i.e., one easy, one medium, and one difficult item), were administered to participants. Additionally, prior to completing each scale, participants were presented with a sample item for the scale. Performance on the TORR has previously demonstrated stability through test-retest reliability (Alexander et al., 2015), and in the present study total TORR scores during Study 1 were highly correlated with scores on the shortened version of the TORR during Study 2 (r = .73, p < .001), indicating good test-retest reliability in the present study.

5.1.3. Study 2 procedures
All data-collection for Study 2 was completed at least two weeks after data collection for Study 1. This two-week period of wait-time is commonly observed in studies using a test-retest methodology, and has been determined to be acceptable for re-testing measures of cognitive ability (Arthur & Day, 1994; Collie, Maruff, Darby, & McStephen, 2003). This length of time provides sufficient time to reduce carry-over effects and memory for item-specific details. While participants reasoned with the TORR items, they were instructed to think aloud. Specifically, they were directed to verbalize anything they were thinking or doing while completing the problems—a procedure consistent with traditional protocol analysis (Ericsson & Simon, 1984; Pressley & Afflerbach, 1995). If participants were silent for more than 15 s, they were reminded to continue thinking aloud. Prior to the study, participants were given the opportunity to practice thinking aloud with two mathematics problems in order to familiarize them with the procedure. The coding scheme is described in detail following a description of the analytical methods.

5.1.4. Bayesian network analysis
The research questions in this investigation specifically pertained to differences among groups at different phases of the reasoning process. Therefore, in order to address this question, Bayesian network analysis was conducted to calculate the conditional probabilities that a participant would reach each phase of the reasoning process given that they had reached the previous phase. Bayes’s theorem, a general formula for conditional probabilities is:

\[ P(A|B) = \frac{P(A) \times P(B|A)}{P(B)} \]

In the present analysis, \( B \) refers to the preceding component process, and \( A \) refers to the subsequent component process, and \( P(A|B) \) can be read as the probability of \( A \) given \( B \). In the equations and analyses that follow, the component processes of encode (E), infer (I), map (M), and apply (A) will be abbreviated using the first letter of the process, such that \( P(\text{infer}|\text{encode}) \) is represented as \( P(\text{IE}) \). The following equation represents the likelihood that an individual will reach the infer phase, given that they have encoded:

\[ P(\text{I}|\text{E}) = \frac{P(\text{I}) \times P(\text{E}|\text{I})}{P(\text{E})} \]

Importantly, in order to utilize Bayes’s theorem to calculate \( P(\text{IE}) \), the related conditional probability \( P(\text{E}|\text{I}) \) must be included in the calculation. In Bayesian analysis, probabilities like these can be termed prior likelihoods, and are often estimated based on theoretical or logical assumptions (Neapolitan, 2004). As indicated, logically and by the definitions used in coding these data, \( P(\text{E}|\text{I}) \) must equal 1. This is to say, if participants verbalized an inference about a test item, then by definition they must have encoded at least some of the visual information presented in that test item. In the same way, \( P(\text{I}|\text{M}) \), and \( P(\text{M}|\text{A}) \), must also equal 1 in calculations pertaining to those phases. Therefore, prior conditional probabilities were set to 1 for this investigation, and the resulting equation for the calculation of \( P(\text{I}|\text{E}) \) is pictured below as an example:

\[ P(\text{I}|\text{E}) = \frac{P(\text{I}) \times 1}{P(\text{E})} \]

Conditional probabilities were calculated separately for each of the three processes that had a subsequent process [i.e., \( P(\text{IE}) \), \( P(\text{MI}|\text{I}) \), and \( P(\text{AA}|\text{M}) \)]. This means that differences in the encoding phase could not be compared. Fig. 1 provides a conceptual depiction of the conditional probabilities associated with each component process.

5.1.5. Coding
After data collection was complete, participants’ think aloud protocols were transcribed and coded. Specifically, protocols were coded for evidence of each of the component process [i.e., encode, infer, map, apply], for each participant, codes were given for each of the 12 items (see Table 5 for a description of the coding scheme and sample codes), and each item received four codes representing the presence or absence of each component process. Utterances were coded as encode if participants noted any identifiable features of items (e.g., color, number, shape, or fill). For example, one participant verbalized that a given element of an item was “two dots in a triangle.” Codes of infer were applied when individuals reported making connections between or among encoded features of a problem. This code reflected the formation of lower-order relations among features of an item (Gentner, 1983, 2010). For example, one participant verbalized an inference when she said “…so the outlines of each shape kind of inverse. The inside shape goes to the outside.” Although it was possible for participants to articulate multiple instances of encoding and inferencing, items were only coded once for whether or not participants engaged to some extent in each process.

Inferences or lower-order relations are not generally considered to be sufficient for relational reasoning, which is described as a higher-order relation mapped among lower-order relations (Crone & Ormerod, 2005). An antithesis or inconsistency in the map of the relation is represented as a conflict, and these codes are termed antitheses (Crone & Ormerod, 2005). For example, one participant verbalized a conflict when she said “I think that’s a triangle…” The衄al figure has been adapted from that study. A single item was coded as an antithesis each time a conflict was noted by a participant.

Table 5
Comparison of low- and high-performing groups’ mean scores on the TORR.

<table>
<thead>
<tr>
<th></th>
<th>Low performers n = 16</th>
<th>High performers n = 16</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>All items</td>
<td>5.56 (SD = 2.37)</td>
<td>9.50 (SD = 1.41)</td>
<td>5.72</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Analogy</td>
<td>1.38 (SD = 0.81)</td>
<td>2.19 (SD = .83)</td>
<td>2.82</td>
<td>.009</td>
</tr>
<tr>
<td>Anomaly</td>
<td>1.63 (SD = .96)</td>
<td>2.63 (SD = .62)</td>
<td>3.51</td>
<td>.002</td>
</tr>
<tr>
<td>Antimony</td>
<td>1.31 (SD = .95)</td>
<td>2.25 (SD = .58)</td>
<td>3.38</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Antithesis</td>
<td>1.25 (SD = .77)</td>
<td>2.44 (SD = .73)</td>
<td>4.47</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Note: Total score for all items out of 12; Scale scores out of 3.
we can say that if a participant successfully applied, they must have mapped the relation. Further, they must have previously formed the necessary lower-order inferences among which the higher-order pattern was mapped, and before that must have encoded relevant features of the problem that allowed for the formation of inferences. Therefore, even if a participant did not verbalize a specific component process in their think aloud, but reached a subsequent process, the item was coded as reaching all previous processes. This logical and theoretical definition of the relational reasoning process informed the Bayesian analysis we utilized in this investigation.

5.1.6. Interrater agreement
For the purpose of the present study, it was important to establish that the four component processes commonly associated with analogy (Sternberg, 1977) were, in fact, present across all relational reasoning forms. Moreover, it was essential to examine whether additional component processes were necessary to describe the process of relational reasoning in any of the four forms. Coding was undertaken by the first and second authors, who independently coded 22.2% of the items for the presence of the targeted processes (e.g., infer), while remaining alert to the verbalization of any additional component processes associated with reasoning performance. During the coding process, several strategies such as guessing and response elimination were identified across each of the four relational forms. However, these strategies represented domain-general test-taking strategies rather than processes specific to relational reasoning (Ellis & Ryan, 2003; Hong, Sas, & Sas, 2006). A high level (87.5%) of exact agreement was found ($\kappa = .83$), and all cases of disagreement were resolved through discussion. To verify the coding scheme, a research assistant who was not associated with the project and was not informed of any of the conditions, research questions, or hypotheses coded 21.9% of the data. The interrater agreement was high (84.5% exact agreement; $\kappa = .79$) and the second author coded the remainder of the data. Data were analyzed using Stata 13 and Cramer’s V effect sizes are reported throughout to identify small ($V < .07$), medium ($V = .21- .35$), and large ($V > .35$) effects (Sun, Pan, & Wang, 2010).

5.2. Study 2 results
To answer RQ2.1, conditional probabilities for each phase of the relational reasoning process were calculated among all students across all items in order to examine overall differences in reaching each processing phase. Next, to address research questions 2.2, 2.3 and 2.4, conditional probabilities were compared for three different groupings: (a) high- and low-performing participants (RQ2.2), (b) easy, medium, and difficult items (RQ2.3) and (c) analogy, anomaly, and antithesis items (RQ2.4). Finally, the role of VS-WMC at each phase of the reasoning process was considered (RQ2.5). Given the potential concerns related to multiple comparisons, a

<table>
<thead>
<tr>
<th>Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding scheme for component processes in think aloud protocols.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Code</th>
<th>Explanation</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encode</td>
<td>Identifying attributes of any parts of the items in the problem</td>
<td>“two dots in a triangle” “this has all these cross-looking things”</td>
</tr>
<tr>
<td>Infer</td>
<td>Identifying a relation (e.g., similarities, differences) between two items; this is the lower-order relation</td>
<td>“A doesn’t seem to have anything that has in common [with the given]” “… the bottom row has a dot on both the left and right side, but the rest aren’t that way as far as the rows are concerned. They don’t match.”</td>
</tr>
<tr>
<td>Map</td>
<td>Identifying a relation between more than one relation; this is the higher-order relation</td>
<td>So, first row has one dot in each, second row has two dots in each shape, so the third row is going to have three dots in each shape. “They all seem to be split directly down the middle … except for C” “So that would be the one with the grid, the black square, and the white triangle”</td>
</tr>
<tr>
<td>Apply</td>
<td>Determining the response that completes the relation based on identified relations/rules</td>
<td>“which is D”</td>
</tr>
</tbody>
</table>
Bonferroni correction was applied to maintain a family-wise error rate of $\alpha = .05$ for all sets of simultaneous tests. Specifically, the significance threshold was set at $\alpha = .05/3$, or $\alpha = .02$, in order to account for sets of three simultaneous tests across the conditional probability levels $P(I|A)$, $P(M|I)$, $P(A|M)$ for each of the research questions.

### 5.2.1. Transitions across component processes

Research question 2.1 addressed whether the conditional probabilities [i.e., $P(I|E)$, $P(M|I)$, $P(A|M)$] were the same for each process. To examine this, the data from the high- and low-performing students were aggregated across items. The conditional probabilities were $P(I|E) = .88$, $P(M|I) = .72$, and $P(A|M) = .93$. An omnibus chi-square test indicated that these probabilities were significantly different [$\chi^2 = 52.25$, $df = 2$, $p < .001$, $V = .23$]. As Fig. 2 suggests, there was a drop in $P(M|I)$ compared to the other two calculated conditional probabilities. This implies that the mapping phase of the relational process was significantly more difficult (as evidenced by the lower probability of successfully reaching that phase) for the combined sample of high- and low-performing participants than the encode, infer, or apply phases.

### 5.2.2. High and low performers

Next, we addressed research question 2.2, that is, do individuals identified as high performing and low performing differ in terms of the ease with which they reached each component processes? To address this question, the conditional probabilities at each phase of the reasoning process were calculated separately for the high and low performers (see Table 6). Chi-square tests were used to confirm that these probabilities were significantly different across the four reasoning processes for both the low performing [$\chi^2 = 36.05$, $df = 2$, $p < .001$, $V = .29$] and high performing [$\chi^2 = 18.51$, $df = 2$, $p < .001$, $V = .19$] participants. Similar to the full sample, the probabilities indicated that the mapping phase was more difficult for both the low and high performers than the other phases in reasoning process (see Fig. 2). Comparing across high- and low-performing students, the conditional probabilities significantly differed at the inference [$\chi^2 = 33.89$, $df = 1$, $p < .001$, $V = .21$] and mapping [$\chi^2 = 30.84$, $df = 1$, $p < .001$, $V = .30$] phases, but not at the application phase [$\chi^2 = .62$, $df = 1$, $p = .43$, $V = .05$]. Examination of the probabilities suggests that the increase in difficulty at the inference and mapping phase was greater for the low-performing students. This implies that the inference phase and especially the mapping phase represented problematic junctures in the reasoning process for low-performing students. However, low- and high-performing students did not differ in their ability to apply a mapped pattern, provided that mapping had occurred.

### 5.2.3. Item difficulty

Research question 2.3 asked whether the conditional probabilities for each phase of the reasoning process differed according to the difficulty of the items. To answer this question, items were divided based on their empirically observed difficulty level. Conditional probabilities for each phase of the reasoning process were calculated separately for the easy, medium, and difficult items (see Table 6). For the full sample of participants, chi-square tests indicated significant differences in the conditional probabilities across the reasoning process for the easy [$\chi^2 = 11.36$, $df = 2$, $p < .001$, $V = .20$], medium [$\chi^2 = 13.71$, $df = 2$, $p < .001$, $V = .20$] and difficult [$\chi^2 = 29.46$, $df = 2$, $p < .001$, $V = .32$] items. Importantly, the conditional probabilities at the inference phase [$\chi^2 = 3.67$, $df = 2$, $p = .16$] did not differ significantly across item difficulty level. However, $P(M|I)$ did significantly differ across the easy, medium, and difficult items [$\chi^2 = 28.06$, $df = 2$, $p < .001$, $V = .28$], as did $P(A|M)$ [$\chi^2 = 9.21$, $df = 2$, $p < .001$, $V = .20$]. This finding implies that the principle difference between easy, medium, and difficult relational reasoning items may have been the complexity of the higher-order pattern patterns were required to map in order to correctly respond to the item and the challenge of applying the mapping for more difficult items. This finding further demonstrates that the mapping phase of the relational reasoning process was the critical phase in determining students’ ultimate reasoning performance.

### 5.2.4. Forms of reasoning

Next, to address research question 2.4, conditional probabilities associated with each phase of the reasoning process were compared across each form of relational reasoning included on the TORR (i.e., analogy, anomaly, antinomy, antithesis). A recent investigation of the factor structure of the TORR (Alexander et al., 2015) revealed that although the scales of the TORR are related to each other they appear to represent separable yet related forms of reasoning. Therefore, differences in the reasoning process among the scales of the TORR are potentially interesting to investigate. Importantly, a within-subjects ANOVA was used to confirm that mean performance across all participants in Study 2 did not significantly differ among the shortened scales of the TORR [$F(3,124) = 1.02$, $p = .394$], implying that the shortened scales of the TORR used in Study 2 did not systematically differ in difficulty. Although item difficulty did not differ across the forms, how students proceeded through the process of reasoning of the different

| Participants | Items | $P(I|E)$ | $P(M|I)$ | $P(A|M)$ |
|--------------|-------|---------|---------|---------|
| Full sample  | All Items | .876 | .724 | .933 |
| Easy         | .928 | .872 | .99 |
| Medium       | .874 | .720 | .90 |
| Hard         | .834 | .566 | .883 |
| Low performers | All Items | .777 | .571 | .916 |
| Easy         | .854 | .773 | .975 |
| Medium       | .777 | .551 | .814 |
| Hard         | .714 | .355 | .721 |
| High performers | All Items | .973 | .844 | .943 |
| Easy         | 1.00  | .953 | 1.00 |
| Medium       | .968 | .854 | .943 |
| Hard         | .953 | .721 | .863 |
forms, and whether certain processing phases proved more challenging in particular forms were examined (see Table 7 for specific calculated conditional probability values).

To understand the degree of similarity or distinctness in the component processes across form, conditional probabilities for each of the phases were compared across forms. Interestingly, significant differences among the reasoning forms were observed for $P(M|I)$ [$\chi^2 = 9.41, df = 3, p = .02, V = .07$] and $P(A|M)$ [$\chi^2 = 9.73, df = 3, p = .02, V = .20$, but not at $P(I|E)$ [$\chi^2 = 6.87, df = 3, p = .08, V = .13$]. This implies that although the forms of reasoning may differ in terms of mapping and application, they appear to be similar in terms of making inferences. As Fig. 3 illustrates, the probability of encoding some visual information from an item was uniformly high across the scales of the TORR. Values for $P(I|E)$ began to decline, showing that the inference phase of reasoning was more difficult for students than the encoding phase. However, the non-significant difference at the inference phase implies that although inferences were challenging, this component process was similarly difficult across forms of reasoning.

At the mapping phase, some potentially important patterns arose. Perhaps foremost among them is the observation that, for difficult inferences were challenging, this component process was similarly significant to decline, showing that the inference phase of reasoning was more difficult for students than the encoding phase. However, the non-significant difference at the inference phase implies that although inferences were challenging, this component process was similarly difficult across forms of reasoning.

5.2.5. Working memory capacity in component processes

Next, analysis was undertaken to determine whether the difficulty of each phase of the relational reasoning process was related to VS-WMC (RQ2.5). Specifically, correlations were calculated between participants’ VS-WMC (calculated using principal component analysis from Study 1) and the conditional probability that participants reached each component process (see Fig. 4). The correlation between VS-WMC and the probability a participant encoded any visual information of a given item was $r = .02, p = .92$, implying that this phase of the reasoning process was not reliant on working memory. Similarly, $P(A|M)$ [$r = .36, p = .04$] was not significant at the corrected alpha level of .02. However, significant correlations were found between VS-WMC and $P(I|E)$ [$r = .43, p = .014$] and $P(M|I)$ [$r = .57, p < .001$]. This pattern of correlations implies that working memory was important for identifying both lower- and higher-order relations. Moreover, as Fig. 4 illustrates, the strength of the relation differed by component process. Specifically, the inference process appeared to be moderately related to working memory, while the mapping phase has a stronger relation. This suggests that mapping was the reasoning component most contingent upon working memory resources. This finding is consistent with previous theoretical accounts and empirical findings (Cho et al., 2007; Krawczyk et al., 2008; Richland & Burchinal, 2013) of relational reasoning and working memory, in which WMC is deeply implicated in the ability to integrate lower-order inferences into higher-order patterns.

5.3. Study 2 discussion

Study 2 provided a probability-based analysis to understand how students progress through the component processes and uncover the phases that proved to be stumbling blocks. For both high- and low-performing students, reaching the mapping phase was more challenging than the preceding phase, even for those students who had successfully negotiated those earlier junctures. In effect, if students reached the mapping phase, they were fairly
successful in applying the mapping to determine the answer. This was similar across items at all difficulty levels and across the four forms of relations examined in the present study, with certain exceptions. Specifically, application was the most challenging component process for the antithetical relations, and the decline in mapping was more prominent for difficult items.

Of particular importance in Study 2 were the identified differences in high- and low-performing students across the component processes. Although these students differed significantly in their ability to correctly solve relational reasoning problems, they were equally likely to determine the correct answer if they had mapped. Yet, for low-performing students, mapping represented a substantial challenge in the process of relational reasoning, particularly when compared to their higher-performing peers. Finally, consistent with Study 1, working memory was a significant predictor of relational reasoning. However, Study 2 parsed this relation at each of the component processes to uncover the relative importance of working memory for certain components compared to others. Specifically, the role of working memory increased across the inference and mapping phases before dropping off at the application phase. This suggests that working memory has an important influence as students determine lower-order relations and an even greater influence as they determine higher-order relations. In this way, different requirements of working memory at the different phases of relational reasoning may account for observed challenges in the process for low-performing students.

6. Conclusions and implications

In this investigation, we delved into the performance (Study 1) and process (Study 2) of relational reasoning. Findings from these studies clarified the role of individual differences across multiple forms of relational reasoning, and revealed that the process of relational reasoning, particularly inferring and mapping, is impacted by individual differences, item characteristics, and relational form.

6.1. Clarifying the role of individual differences in multiple forms of relational reasoning

Given the demonstrated importance of relational reasoning for learning and academic achievement (e.g., Farrington-Flint et al., 2007; Richland & McDonough, 2010; Schiff & Ravid, 2007), what we wanted to uncover from Study 1 was whether select individual difference variables were associated with students’ performance of relational reasoning across multiple forms of relations. What we uncovered in Study 1 was that all forms of relational reasoning behaved similarly to other demanding cognitive measures (e.g., Raven's matrices, mental rotation) when it came to visuospatial working memory (Carpenter et al., 1990; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). In effect, successful performance on a measure such as the TORR, rests to some extent on individuals’ capacity to retain symbolic information in working memory. Then, by breaking down relational reasoning into its component processes, Study 2 revealed that VS-WMC was important to the performance of relational reasoning due to its role in making inferences and mapping patterns.

In Study 1 we also determined that gender mattered for TORR performance in that males outperformed females on this figural reasoning task. Importantly, however, that effect dissipated when visuospatial working memory was controlled, and this difference only occurred for the analogical reasoning scale and the overall TORR score. This suggests that certain individual differences such as gender may not play the same role in all forms of relational reasoning. Further, we included need for cognition in Study 1 with the expectation that this motivational factor, which had proven influential in some studies of analogical reasoning (Day et al., 2007; Fleischhauer et al., 2010; Hill et al., 2013), would also be a significant contributor to the forms of relational reasoning included in this study. Theoretically, it is believed that effortful processing should lead to improved reasoning. Yet, that expectation was not met, as the present study did not find a significant correlation between NFC and any of the forms of relational reasoning. One explanation for this lack of an effect turns on the rather novel but challenging character of the TORR, which may have leveled the field and overridden the effect of individuals’ general predisposition to engage in and enjoy thinking and cognitive tasks.

6.2. Explicating the process of relational reasoning

Beyond clarifying the contributors to relational reasoning performance (Study 1), we wanted to understand the underlying cognitive components that may have facilitated or frustrated performance (Study 2). In other words, we sought to understand the fundamental processes that serve as the basis for relational reasoning success in classrooms. For instance, when biology students are unsuccessful in their attempt to draw an analogy between a cell and a factory, what points in this process may serve as roadblocks inhibiting their success? And, is working memory equally important across all component processes as psychology students attempt to determine whether the theories of Piaget and Vygotsky are mutually exclusive? Using probability-based analyses, we wanted to determine whether underlying componential processes (Sternberg, 1977) manifested differently as a consequence of the overall capability of individuals, the form of relational reasoning, and item difficulty.

6.2.1. Componential analysis by reasoning form

All four forms of relational reasoning involved the same component processes, and across all forms, students were almost universally successful with encoding, the most basic of the component processes. However, our probability-based analyses revealed some intriguing differences that merit consideration. Specifically, it was the inferring to mapping phase that was more challenging for anomaly items, whereas the mapping to application phase was more demanding for respondents solving the antithesis problems. What these data suggest is that there were fluctuations or iterations in the difficulty of the processes beyond the encoding phase that must be considered when seeking to discern patterns of similarity, aberrance, incongruence, or opposition.

6.2.2. Processing patterns for high- and low-performing students

As Sternberg and colleagues (Sternberg, 1977; Sternberg & Rifkin, 1979) have demonstrated for analogical reasoning, this study revealed that the mapping phase proved the most demanding component for all performance levels. That is understandable because it is when mapping occurred that individuals made the pivotal higher-order relation that ultimately defines the particular relational pattern. In effect, even if respondents were successful in attending to or perceiving features within the problem space (encode) and making local or low-level associations (infer), there was no guarantee that they would ultimately reach understanding of the higher-order similarity, aberrance, incongruence, or opposition relation that gave patterns their definitive character.

Indeed, when we looked at processing for items that were relatively easy, moderately difficult, and difficult, we found that it was the component of mapping that accounted largely for item difficulty—a pattern that was true for both low performers and high performers. Yet, low and high performers could also be distinguished by their ability to infer. That means that some of the
barriers to successful processing for low performers occurred prior
to mapping, when they attempted to build local associations within
the information stream.

6.3. The integration of individual differences and item
characteristics in the process of relational reasoning

To understand the major stumbling blocks for low performing
students, it is critical to examine how the factors of item difficulty,
item type, and individual characteristics such as working memory
collectively impact student performance. In other words, is it the
characteristics of the students, the items, or a combination that best
explain why some students struggle to create inferences and map
higher-order relations? Results from each of the research questions
in Study 2 point to inferring and mapping as the main bottle-necks
in the reasoning process for low-performing students. In particular,
when items are difficult, low-performing students experience
additional challenges when making inferences and connecting
those inferences to map an overarching relation. This observation is
consistent with theories of relational complexity, which suggest
that relational reasoning is more difficult when items require stu-
dents to combine more inferences in creating a higher-order rela-
tion (Birney & Halford, 2002; Cho et al., 2007). Findings from Study
1 provided evidence that low-performing students had lower VS-
WMC on average, a characteristic that Study 2 found to be highly
related to whether students could infer given that they had enco-
ded, or map given that they had inferred.

6.4. Limitations and future directions

As with any study, it is important to consider these conclusions
in concert with the limitations of the study, and to use the con-
cclusions and limitations to determine future directions. First,
although the present study extended research on the processes of
analogical reasoning to include multiple forms of relational
reasoning, the additional forms tested in the present study (i.e.,
anomaly, antinomy, antithesis) do not necessarily encompass all
possible relational forms. The present studies and previous theo-
retical arguments (Alexander & the DRLR, 2012; Dumas et al.,
2013) do not claim that these are the only forms of relational
reasoning. Therefore, future research is needed to examine
whether additional relational forms exist and are consistent with
the present findings. Moreover, given that the TORR was the only
existing measure of relational reasoning that included multiple
forms of reasoning, the questions in Study 2 were drawn from the
TORR items in Study 1. To alleviate test–retest effects, adequate time
was provided between testing sessions, resulting in expected dif-
fferences between high- and low-performing students. However, it
is possible that individual students switched from high to low (or
low to high) performers from Study 1 to Study 2. Despite the po-
tential for some individual shifts in performance, there was strong
stability in TORR performance from Study 1 to Study 2, and sig-
nificant differences between high and low performers were present
for all subscales of the TORR. The availability of other multi–form
measures of relational reasoning will help to alleviate possible
testing effects in future studies.

Second, in order to access students’ underlying on-line pro-
cesses, Study 2 utilized a think aloud methodology. The decision to
use think aloud methodologies was based on our interest in stu-
dents’ articulation of their conscious processes, and the precedence
of this methodology for uncovering processing differences across a
range of ability levels (Ericsson & Simon, 1980; Pressley &
Afflerbach, 1995; Wineburg, 1991). Yet, it is important to
acknowledge that, with any method, think alouds have certain
shortcomings, which have been reviewed elsewhere (e.g., Ericsson
& Simon, 1984). In particular, some students may be better at
articulating their processes than other students, resulting in an
underestimate of instances of the relational processes for certain
students. Alternatively, some instances of pattern application might
have been overestimated if the participant guessed correctly but
did not explicitly state that they were guessing. Additionally, the
time-consuming nature of think alouds often precludes the use of a
large sample or a large number of items for students. In Study 2, a
restricted number of students and items allowed for the collection
of this data; however, it limited our ability to capture small effects.
Think alouds are a beneficial tool for uncovering conscious pro-
cesses; however, particularly in expertise, automatic processes are
often masked through methodology because the processes are
believed to occur below the level of conscious awareness (Ericsson
& Simon, 1980). Moreover, given that think alouds have been criti-
cized for their potential to influence on-line processes (DeShon,
Chan, & Weissbein, 1995; Lane & Schooier, 2004), future research
should examine methods such as brain imaging and eye-tracking,
which allow for the examination of on-line processing without
requiring participants’ verbal reports.

Third, when identifying instances of encoding, inferring, and
mapping, we were able to code for the presence of any instances,
but were not able to code for the accuracy or relevance of such
instances. In other words, participants may have identified an
incorrect inference resulting in an incorrect mapping, but would
still have received credit for engaging in those processes. Moreover,
participants received credit for the application stage if they
responded correctly without guessing. In order to verify the accu-
racies of participants’ processes, it is necessary to know what specific
part of the problem the participant is viewing. This may be possible
for future studies using concurrent eye-tracking and verbal report
methodologies, and would provide important information beyond
what was available in the present study.

Fourth, given time constraints for participation, we were limited
in our examination of individual characteristics in the present
study. For example, motivation is typically regarded as a multifac-
et construct that iteratively interacts with cognitive processing.
In Study 1, NFC was selected for its potential relevance in explaining
relational reasoning, however it represents only one form of
motivation. Future research should consider the role of domain-
and task-specific motivation in relational reasoning. Additionally,
we were not able to examine interactive effects between individual
differences and relational reasoning. For instance, NFC in the pre-
sent study was only examined as a predictor of reasoning, not a
consequence. Although NFC did not predict relational reasoning in
the present study, it is important to examine whether the relation
between these variables is better understood as an iterative pro-
cess. Future studies that examine the reciprocal relation between
motivation and relational reasoning will provide important infor-
mation about how motivation influences and is influenced by
reasoning.

Finally, our examination of WMC suggested that a higher
working memory capacity is positively related with relational
reasoning performance. Future research should examine which
specific working memory components (e.g., central executive
components) are most critical for success in the process of rela-
tional reasoning, and whether this differs for high- and low-
performing students. In addition, alternative hypotheses should
be noted. For instance, it is important to consider whether in-
dividuals who perform better on relational reasoning tasks have
the ability to create more advanced mental models, therefore
requiring less working memory and allowing them to reason more
effectively than lower-performing students. Further, although cor-
relations between relational reasoning and gender differences were
identified, these need to be interpreted with caution given the
relatively small sample in the present study. Future investigations should study gender differences with larger and more diverse samples. Based on these data, it remains unclear whether the observed gender differences would be present in a sample of natural or physical science students, who may receive more explicit training and practice on visuospatial rotation than the social science sample in the present study.

7. Concluding thought

Without the ability to discern meaningful patterns within the interminable flood of information that is a hallmark of 21st century, individuals would remain seriously hampered in their ability to survive and to thrive. Deconstructing relational reasoning is not only of theoretical relevance but can also opened the door to outcomes of practical importance. For instance, it is conceivable that the component processes used to train analogical reasoning (Alexander et al., 1987; White & Alexander, 1986) could be effectively applied to the training of alternative forms of relational reasoning, acknowledging that certain forms might require additional support for particular processes. Thus, the more researchers come to know about relational reasoning, including the cognitive and non-cognitive variables that exert influence and the component processes that undergird more or less successful performance, the more that could be done to foster the foundational ability in individuals of differing backgrounds, ages, and abilities. While there is much more to be learned about the nature and manifestation of relational reasoning, we feel strong that this investigation makes important strides in that direction.

Appendix A. Test of Relational Reasoning (TORR) Sample Items

Analogy

Directions: Below is a pattern that is not yet complete. Select the figure from those shown below that completes the pattern.

Sample Problem 1

A  B  C  D

Triangle  Triangle  Circle  Circle

???
Anomaly

Directions: All of these figures but one follow a particular pattern or rule. Find the one figure that does not follow the pattern.

Sample Problem 1

A

B

C

D

Antinomy

- The problems in this section ask you to compare sets of objects that vary in certain features.
- Each set has a specific rule that decides what objects can be included in that set. Some of the objects included in each set are pictured, enough to allow you to determine its rule for inclusion.
- Every problem asks you to identify which ONE of the four sets that are shown could NEVER have an object in common with the Given set, based on the compatibility of their rules for inclusion.
- There will always be EXACTLY ONE set that is incompatible with the Given set.

Sample Problem 1

GIVEN

A

B

C

D
Antithesis

References


