

# Hidden Structure: Indirect Measurement of Relational Representation

Samuel B. Day (s-day2@northwestern.edu)

Dedre Gentner (gentner@northwestern.edu)

Department of Psychology, Northwestern University  
2029 N. Sheridan Drive Evanston, IL 60208-2710 USA

## Abstract

Structured mental representations underlie much of human cognitive ability. However, research has repeatedly found that people are generally quite poor at spontaneously applying structure acquired from one kind of situation to a dissimilar domain. This would seem to present a serious impediment to learning abstract structures experientially. The current study employs a novel approach designed to be more sensitive to the relational learning that is taking place. By using implicit measures of processing fluency, we demonstrate that relational commonalities—even between dissimilar domains—can indeed exert a measurable, if subtle, influence on subsequent processing. Despite having no explicit awareness of the structural commonalities, participants in our study rated a scientific passage to be better understood, better written, and more interesting when it was preceded by an (overtly dissimilar) analogous passage. This finding has important implications for the acquisition of complex knowledge structures.

**Keywords:** analogy, fluency, relational representation, implicit learning

## Introduction

Human cognition is characterized by a remarkable capacity for acquiring and using complex knowledge structures. Most knowledge involves much more than just a list of independent parts—it contains considerable information about the role that each concept or entity plays relative to the others. In capturing these semantic relationships between elements, structured representations are crucial to human thought—they underlie our ability to generate new inferences, to make accurate predictions, and to transfer existing knowledge to novel problems or domains (see Fodor and Pylyshyn, 1988; Gentner, 1983; Markman, 1999).

It is therefore surprising that participants in many laboratory tasks have shown great difficulty with relational knowledge. The classic example of such difficulty is the research by Gick and Holyoak (1980, 1983) on spontaneous analogical transfer. In those studies, participants were exposed to an instantiation of a particular relational structure, such as a “convergence” structure in which several small forces acted upon a single target to produce a significant effect. Participants were then given an analogous problem from a different domain—that is, a situation with a similar underlying relational structure, but different concrete and contextual details. When participants were made aware of the relationship between the scenarios, they readily used

the prior structure to solve the problem. In the absence of such an explicit hint, however, those studies (and countless replications) found that individuals were very unlikely to use the structure from the prior story to solve the problem, or even to notice the structural commonalities between the two. Additional research has confirmed that while concrete features serve as robust reminders for previous episodes, purely relational knowledge is typically quite poor in this regard, even when the prior structure is demonstrably retained in memory (e.g., Gentner, Rattermann & Forbus, 1993; Ross, 1987). This pattern suggests that the representation of relational knowledge—or at least our ability to access these representations—is both fragile and fleeting.

These findings reflect the vexing problem that educators call “inert knowledge” (Whitehead, 1929), or the inability of students to apply potentially useful knowledge in new contexts. For cognitive theory, however, they pose a mystery and a challenge. It seems intuitively clear that people can acquire some relational knowledge through experiential learning. But if the representation of relational structure is as ephemeral and elusive as the analogical transfer literature suggests, the exact means of this acquisition remains a puzzle.

In the current research, we focus on the early stages of relational learning, and explore the conditions under which incipient representations may become the stronger kinds of knowledge that ultimately drive behavior. In order to detect subtle early representational changes, however, we must employ measures that are more sensitive than those that have generally been used in this area. In this paper, we utilize measures that explore changes in *processing fluency*—the ease with which a stimulus is processed. Some of these measures are obvious, such as the fact that easier processing generally leads to faster responses. Others, such as the fluency-based misattributions discussed below, are much less intuitive.

Drawing on this literature, we examine whether processing a relational structure instantiated in one context facilitates its processing in a new context. Specifically, we use these measures to explore whether a target passage will be processed more easily after exposure to an analogous passage, even in the absence of conscious awareness of this analogy. We also explore the role of representational strength in these effects by examining the impact of prior knowledge or expertise on these processing changes. First,

we briefly review the literature on processing fluency, and some related research on relational learning.

### Processing fluency

It is well-established that for most types of stimuli, repeated exposure leads to facilitated processing—this fact underlies basic repetition priming. Recently, a considerable amount of research has expanded upon this phenomenon by examining situations in which people use their own ease of processing, or “fluency”, as a source of information in itself (see Schwarz, 2004). Specifically, it has been shown that the recognition that a stimulus is easy to process may lead individuals to make inferences and attributions about the cause of this facilitation.

For example, Witherspoon and Allan (1985) asked participants to judge the duration of words that were flashed on a computer screen. They found that these estimates were significantly longer for words that had been seen in an earlier phase of the experiment. The authors argued that the prior exposure had made these words easier to process, but that the participants were misattributing this fluency to the salient dimension of presentation duration—that is, they incorrectly inferred that the words were easy to process *because* they were presented for longer. This sort of misattribution has been shown across a range of dimensions—as a striking example, Mandler and colleagues (Mandler, Nakamura, & Van Zandt, 1987) found that participants could be led to indicate that a fluently processed stimulus was either lighter *or* darker than a control stimulus, depending on which question they were asked. It seems that individuals are in some way able to recognize that a stimulus is easy to process, but that in the absence of knowledge about the true source of this fluency they are willing to attribute the facilitation to a variety of sources—frequently whatever dimension is made salient by the current task.

This effect has been shown to operate in the opposite direction as well: increasing processing fluency by perceptually manipulating a stimulus often leads to higher rates of recognition (including false recognition). For instance, participants are more likely to endorse a word as previously seen when it is presented with a light rather than a heavy visual mask at test (Whittlesea, Jacoby, and Girard, 1990).

While these fluency-based attributions generally seem to take place quickly and effortlessly, they are not simply perceptual effects—they have been shown to depend critically on an individual’s beliefs and expectations (e.g., Schwarz, 2004; Unkelbach, 2007; Westerman, Lloyd & Miller, 2002; Whittlesea & Williams, 2001a&b).

### Structural priming

Several researchers have used implicit measures to explore other forms of structure, such as perceptual or syntactic structure (e.g., Bock, 1986; Gordon & Holyoak, 1983; Reber, 1967). For example, Gordon & Holyoak (1983) reported a “structural mere exposure effect,” in

which exposure to character strings conforming to a particular abstract “grammar” led to higher liking ratings for new grammatical strings. Other researchers have looked at priming for individual semantic relations, such “lives in” (e.g., BIRD–NEST), with somewhat mixed results (Estes & Jones, 2006; Gagne, Spalding & Ji, 2005; McKoon & Ratcliff, 1995; Spellman, Holyoak & Morrison, 2001).

The current research differs from these previous efforts not only in the nature of the measurements used, but also in the use of large-scale semantic structures, involving systems of semantic relations rather than individual relational concepts. A concomitant difference is a much longer delay between processing a relation in the base and target than that used in prior studies.

### Experiment

The current experiment explores the possibility of increased processing fluency for abstract relational structures. Specifically, we investigate whether a text passage will be processed more easily when it is preceded by an analog (i.e., a passage with the same relational structure but different concrete content) than when it is preceded by a matched control passage. In keeping with previous work in analogy, we refer to the first presentation of a particular instantiation of a structure as the *base*, and the subsequent presentation of the structure instantiated in a new context as the *target*, reflecting the fact that the former may serve as a source of information for the latter.

Given the relative novelty of our approach, we included several different dependent measures of fluent relational processing, including changes in reading time, improvements in comprehension, and misattributions to various aspects of the text.

Processing time is often used as a measure of fluent processing; for example, here it could be predicted that a target passage may be read more quickly when it is preceded by an analog. However, it is important to remember that text passages differ in important ways from many other kinds of stimuli—they are quite complex, and their processing unfolds over an extended period of time, allowing for much more strategic variation. Furthermore, it is possible that increased fluency may actually encourage greater interest and therefore greater elaboration of the material, potentially leading to *longer* reading times. The effects of fluency on reading speeds are therefore an open empirical question.

We also explored three areas in which the fluent processing of a text might lead to misattributions: subjective comprehension of the target passage, interest in the topic of the passage, and the perceived quality of writing. (We also examined the effects of relational fluency on participants’ *actual* comprehension of the target passage, as measured by performance on content questions. However, this measure proved to be problematic, due to participants’ introduction of outside knowledge to the task, and will not be discussed here.)

Finally, we also investigated how these effects may vary with expertise. Specifically, we predicted stronger effects for individuals with relatively little prior knowledge about the target domain. We reasoned that for individuals with well-established domain knowledge, strong existing mental representations would leave less opportunity for any benefits of recent exposure.

The study uses passages from rich, real-world domains; specifically, they draw on an extended analogy between the domains of genetics and computer information storage. These domains were selected both because they represent a good structural match, and because they seem well suited to exploring issues of expertise in fluency, since they are areas in which prior knowledge can vary greatly.

### Participants

One hundred Northwestern undergraduates participated in this study for partial course credit.

### Materials

The materials consisted of expository texts from fairly technical domains, each about half a page in length (with an average of 328 words per passage). In addition to a filler passage describing the operation of lasers, participants read a base or control passage describing computer information storage, and a target passage describing aspects of genetics. The information in the base and target passages drew on a number of structural commonalities between these domains (see Table 1 for a list of correspondences). In order to avoid ceiling effects, the target passage was intentionally written to be somewhat opaque.

The control passage also described computer storage—specifically, data compression. The passage was similar in length and style to the base text, but did not share meaningful structural commonalities with the target. The passages were also controlled for wording, such that the base and control texts did not differ substantially in the content words that each shared with the target. More specifically, there were only seven words uniquely shared by the base and target passages (*also, become, into, it, multiple, one, and sometimes*), and eight words uniquely shared by the control and target passages (*frequently, involves, may, necessary, sequence, string, than, and within*). Furthermore, latent semantic analysis (Landauer & Dumais, 1997) produced similar ratings of semantic overlap between the target and the base and control passages (.25 and .30, respectively). These steps helped to ensure that any observed effects of the base on the target would be solely due to shared abstract structure between the two.

### Design and procedure

The study was implemented as a computer-based task. In the instructions, participants were told that they would be reading passages describing topics in science and technology, and assessing those passages in various ways. They were instructed to read each passage carefully, but not to spend too much time on any one. Participants were not

informed of the additional content questions for the target passage. They then read three passages at their own pace. All participants were first presented with a filler passage describing the operation of lasers. For the second passage, half of the participants (the Analogy group) read the base passage on data storage (which was analogous to the target), while the other half (the Control group) read the control passage describing data compression. Finally, all participants read the target passage on genetics.

After each passage, participants provided various ratings about the text, with each rating entered by clicking on a horizontal bar on the computer screen. The four ratings provided were: (1) Quality of writing: “Please click anywhere on the scale below to rate how clearly the information in the previous passage was presented”, with scale endpoints labeled “Not at all clear, very difficult to understand” and “Very clear and easy to understand”; (2) Interest: “Please indicate how interesting you found the material in the passage”, with endpoints labeled “Not at all interesting” and “Very interesting”; (3) Subjective comprehension: “Please indicate how well you now understand the material that was presented”, with endpoints labeled “Do not understand at all” and “Understand extremely well”; (4) Prior knowledge: “Please rate how much of the material in the passage you already know prior to reading it”, with endpoints labeled “Knew none of it

Table 1: Structural correspondences between passages

Correspondences
• Composed of discrete alphabet (0,1 / A,C,G,T)
• Elements combined into units of determinate size (bytes / codons)
• These units combined into larger meaningful sequences of indeterminate length (files / genes)
• Indicators of starting location for information sequences (file allocation table / promoter site)
• Information sequence is not necessarily stored contiguously
• Information is copied to a new location for use (RAM / RNA)
• During this copying, non-contiguous information is consolidated
• Information about how to use this data is also copied (applications / tRNA, rRNA)
• Original information is not changed during copying
• May be organized such that the same action is performed on multiple units simultaneously
• Some latent information remains stored
• This latent information may be reactivated (via changes to the file allocation table / promoter site)

before reading” and “Knew all of it before reading”. For analysis, the location of the click for each rating was transformed into a continuous value from zero to one.

After rating the final text, participants were presented with questions about the content of the target passage. These were open-ended questions designed to assess comprehension of the relevant structure; for example “How is the organization of the information in a gene different in DNA and RNA?” Participants answered five such questions at their own pace (questions and responses were part of the computer-based task), and were allowed to enter as much or as little information as they wished.

Finally, participants answered a few general questions designed to assess their potential awareness of the structural commonalities between the base and target. They were first asked for their general impression of the task, and whether they had found anything interesting or unusual. Next, they were specifically asked whether they had noticed any commonalities between any of the passages, and if so to describe them. Finally, for those participants who indicated that they had noticed commonalities, they were asked whether these had been noticed during reading or afterwards.

## Results

The purpose of this study was to explore differences in the processing of a target passage resulting from prior exposure to an analogous text, and to see how these differences might vary as a function of prior knowledge. However, initial examination of the data revealed that some participants had read the base or target passage quite quickly. Given that our theoretical questions require adequate encoding of the material, we omitted from analysis any participant who read either the base or target in less than half of the average (mean) reading time for that passage, resulting in the exclusion of 16% of the participants.

Tests revealed that participants in the Analogy condition gave significantly higher ratings for subjective comprehension of the target (.52 v. .38,  $t(1, 83) = 2.94, p < .01$ ), interest (.49 v. .37,  $t(1, 83) = 2.18, p < .05$ ), and quality of writing (.55 v. .44,  $t(1, 83) = 2.15, p < .05$ ), relative to those in the Control condition. Prior exposure to the relevant relational structure thus appears to have facilitated its subsequent processing in a new context. We also verified that, as expected, the Analogy and Control groups did not differ in their ratings of prior knowledge of the target domain, or in any of their ratings for the filler or base passages (all  $ps > .1$ ). Interestingly, there was no significant difference between conditions for the reading time of the target passage (75.4 v. 78.0 sec,  $t < 1$ ).

In order to assess the influence of prior knowledge on these fluency effects, we performed a median split on the ratings of prior knowledge for the target passage, and analyzed the low and high groups separately. In the low prior knowledge group, we again found significant differences on ratings of subjective comprehension (.44 v. .23,  $t(1, 41) = 3.79, p < .001$ ), interest (.43 v. .25,  $t(1, 41) =$

2.84,  $p < .01$ ), and quality of writing (.42 v. .23,  $t(1, 41) = 3.53, p < .01$ ) for the target. However, analysis of the high knowledge group revealed no such differences (all  $t < 1$ ). Thus, although the high knowledge group tended to give higher ratings in general, the experimental manipulation of exposure to an analogy increased the ratings only for those with low prior knowledge of the target domain.

Remarkably, virtually none of the participants indicated that they had noticed the structural commonalities between the base and target passages. When asked at the end of the session (“Did you notice any commonalities between any of the passages, and if so, what were they?”), the most common response was that all of the passages were about science or technology, and several participants simply said no. A few participants in both the analogy (2 Ps) and control (6 Ps) conditions noted that both genetics and computer storage had to do with the storage of information, or that both involved larger components being built from smaller constituents. However, only two of the 50 participants in the analogy condition reported noticing any of the deeper structural commonalities between the domains.

## Discussion

This study provides direct evidence for fluency effects based on structured semantic relations. Participants rated a passage as better written, better understood, and more interesting when it had been preceded by an analogous passage. This occurred even though the passages were from very different domains, and despite the fact that participants uniformly failed to recognize the structural commonalities between the two. These findings are consistent with the hypothesis that participants are forming a representation that includes structural information, and that this representation may facilitate processing for subsequent input that shares this structure.

The study also begins to establish some possible boundary conditions for this type of effect. When people came to the task already in possession of a solid representation for the target material, exposure to the analog appeared to provide

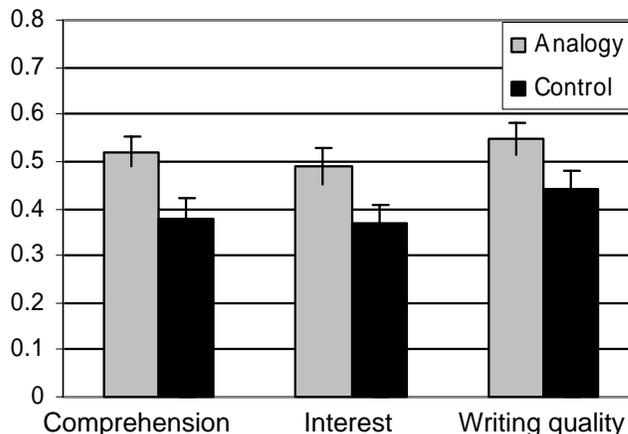


Figure 1: Ratings were significantly higher for participants in the Analogy group

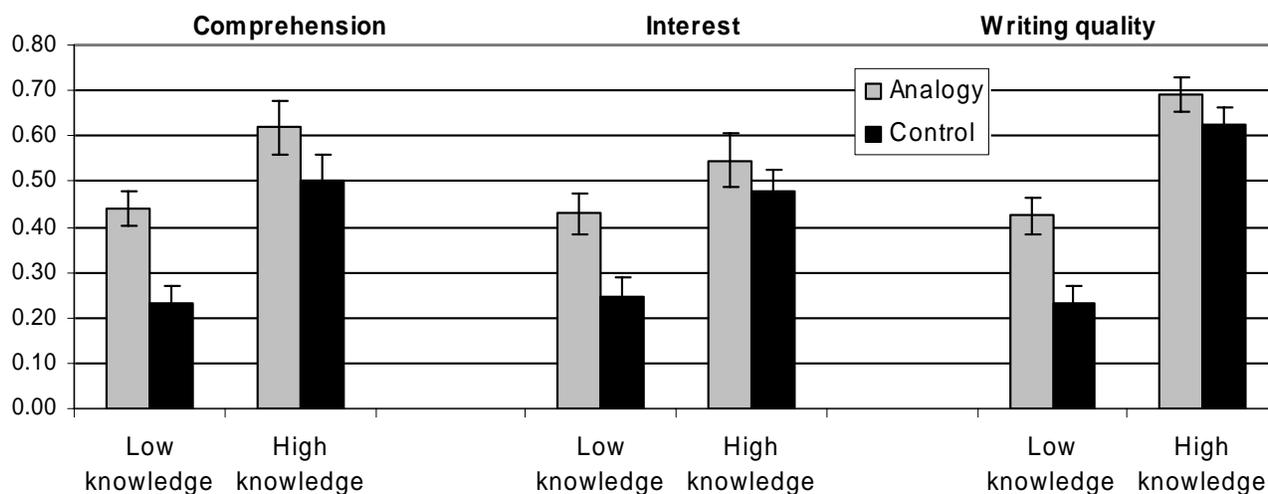


Figure 2: Fluency effects were only seen in participants with low prior knowledge.

no processing benefit. This suggests that structural fluency effects are indeed most appropriate as a measure of the weaker, emerging representations that develop early in the learning process.

Finally, the study provides some insight into appropriate measures of these effects, and into the kinds of attributions that are likely to result from fluent processing of a large semantic structure. Interestingly, fluent processing did not appear to produce faster overall reading times. As noted earlier, individual reading strategies and cognitive styles could have introduced a significant amount of noise to this measure. However, indirect measures of fluency—through subjective ratings—seemed to be a more sensitive indicator of processing changes. To varying degrees, the experimental manipulation increased participants' ratings of comprehension, interest, and writing quality.

### Conclusions

Abstract, structured knowledge representations provide much of the power of human cognition. They allow a person to understand the relationships between concepts and between entities, and to use these relationships to make accurate inferences and predictions across a variety of situations. It has been surprising, then, to find that people seem to have such difficulty in acquiring such knowledge experientially.

The current study provides some important insight into this issue. The results suggest that relational knowledge may be acquired gradually, but that even a single exposure may lead to some degree of representational change. However, these changes may be too subtle to manifest themselves in tasks that require active knowledge application, or even the explicit recognition of common structure. More implicit measures, which explore changes in the simple processing of a stimulus, may be better able to detect the incipient representational changes that are taking place.

In addition to the insights about relational knowledge itself, we believe that this research provides an approach which may open the door for more in-depth explorations of abstract learning. For example, we are beginning to explore how these early representations evolve over time, and how they are affected by the overt similarity between situations. We are also investigating the relationship between these effects and more traditional measures of learning; for example, examining whether early implicit measures are predictive of later performance on learning and transfer tasks. We believe that this work marks a new and fruitful approach to understanding human knowledge.

### Acknowledgements

This work was supported by ONR award N00014-02-1-0078. Thanks to Jennifer Asmuth, and the Northwestern Similarity and Analogy group.

### References

- Bock, J.K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387.
- Estes, Z., & Jones, L.L. (2006). Priming via relational similarity: A copper horse is faster when seen through a glass eye. *Journal of Memory & Language*, 55, 89–101.
- Fodor, J.A., & Pylyshyn, Z.W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3–71.
- Gagne, C.L., Spalding, T.L., & Ji, H. (2005). Re-examining evidence for the use of independent relational representations during conceptual combination. *Journal of Memory & Language*, 53, 445–455.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155–170.
- Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating

- retrievability from inferential soundness. *Cognitive Psychology*, 25, 524–575.
- Gick, M.L., & Holyoak, K.J. (1980). Analogical problem solving. *Cognitive Psychology*, 12, 306–365
- Gick, M.L., & Holyoak, K.J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1–38.
- Gordon, P.C., & Holyoak, K.J. (1983). Implicit learning and generalization of the mere exposure effect. *Journal of Personality & Social Psychology*, 45, 492–500.
- Landauer, T.K., & Dumais, S.T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211–240.
- Mandler, G., Nakamura, Y., & van Zandt, B.J.S. (1987). Nonspecific effects of exposure on stimuli that cannot be recognized. *Journal of Experimental Psychology: Learning Memory & Cognition*, 13, 646–648.
- Markman, A.B. 1999: *Knowledge Representation*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mckoon, G., & Ratcliff, R. (1995). Conceptual combinations and relational contexts in free association and in priming in lexical decision and naming. *Psychonomic Bulletin & Review*, 2, 527–533.
- Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning & Verbal Behavior*, 6, 855–863.
- Ross, B.H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 13(4), 629–639.
- Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, 14, 332–348.
- Spellman, B.A., Holyoak, K.J., & Morrison, R.G. (2001). Analogical priming via semantic relations. *Memory & Cognition*, 29, 383–393.
- Unkelbach, C. (2006). The learned interpretation of cognitive fluency. *Psychological Science*, 17, 339–345.
- Westerman, D.L., Lloyd, M.E., & Miller, J.K. (2002). The attribution of perceptual fluency in recognition memory: The role of expectation. *Journal of Memory & Language*, 47, 607–617.
- Whitehead, A. N. (1929). *The aims of education and other essays*. New York: Macmillan.
- Whittlesea, B.W.A., Jacoby, L.L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory & Language*, 29, 716–732.
- Whittlesea, B.W.A., & Williams, L.D. (2001a). The discrepancy-attribution hypothesis: I. the heuristic basis of feelings of familiarity. *Journal of Experimental Psychology: Learning Memory & Cognition*, 27, 3–13.
- Whittlesea, B.W.A., & Williams, L.D. (2001b). The discrepancy-attribution hypothesis: II. expectation, uncertainty, surprise, and feelings of familiarity. *Journal of Experimental Psychology: Learning Memory & Cognition*, 27, 14–33.
- Witherspoon, D., & Allan, L.G. (1985). The effect of a prior presentation on temporal judgments in a perceptual identification task. *Memory & Cognition*, 13, 101–111.