Ford, M (1995). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition, 54*, 1-71.

Gámez, E., & Marrero, H. (2000). The role of the middle terms contiguity in pragmatic syllogistic reasoning. *Cahiers de Psychologie, 19 (5-6)*, 487-512.

Gingerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650-669.

Glass, A. L., & Holyoak, K. J. (1974). The effect of some and all on reaction time for semantic decisions. *Memory and Cognition, 2(3)*, 436-440.

Holyoak, K. J., & Glass, A. L. (1978). Recognition confusions among quantifiers. *Journal of Verbal Learning & Verbal Behavior, 17(3)*, 249-264.

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic Inference. *Cognition, 16*, 1-62.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hillsdale, NJ: Lawrence Erlbaum.

Manktelow, K., & Over, D. (1995). Deontic reasoning. In S. Newstead & J. St. B. T. Evans (Eds.). *Perspectives on thinking and reasoning: Essays in honour of Peter Wason* (pp. 91-114). Hove, UK: Lawrence Erlbaum.

Moxey, L. M. & Sandford, A. J. (1993). *Communicating quantities*. Hove, UK: Lawrence Erlbaum.

Newstead, S. E. (1995). Gricean implicatures and syllogistic reasoning. *Journal of Memory and Language, 34*, 644-664.

Newstead, S. E., & Coventry, K. R. (2000). The role of context and functionality in the interpretation of quantifiers. *European Journal of Cognitive Psychology, 12(2)*, 243-259.

Oakhill, J. V., Johnson-Laird, P. N., & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition, 31*, 117-140.

Oaksford, M., Roberts, L., & Chater, N. (2002). Relative informativeness of quantifiers used in syllogistic reasoning. *Memory and Cognition, 30(1)*, 138-149.

Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review, 102*, 533-566.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: M.I.T. Press, Bradford Books.

Roberts, M. J., Newstead, S. E., & Griggs, R. A. (2001). Quantifier interpretation and syllogistic reasoning. *Thinking & Reasoning, 7(2)*, 173-204.

Santamaría, C., García-Madruga, J. A., & Carretero, M. (1996). Beyond belief bias: Reasoning from conceptual structures by mental models manipulation. *Memory & Cognition, 24* (2), 250-261.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science, 19*, 97-140.

Stenning, K., & Yule, P. (1997). Image and language in human reasoning: A syllogistic illustration. Cognitive Psychology, *34*, 109-159.

Wheterick, N., & Gilhooly, K. (1990) Syllogistic reasoning: Effects of premise order. In K. Gilhooly, M. T. G. Keane, R. Logie & G. Erdos (Eds.), *Lines of thought: reflections on the psychology of thinking, Vol. 1*. London: Wiley.

Woodworth, R. S., & Sells, S. B. (1935). An atmosphere effect in syllogistic reasoning. *Journal of Experimental Psychology, 18*, 451-60.

Declarative Transfer from a Memory Task to a Problem Solving Task[1]

Timothy J. Nokes and Stellan Ohlsson[2]
University of Illinois at Chicago

Knowledge transfer is the key to understanding the relation between learning and problem solving. In order for declarative knowledge acquired from a learning scenario to transfer to a problem solving situation the knowledge must be both abstract and generative. In a laboratory experiment we examined whether knowledge acquired from a memorization task could facilitate performance on a complex problem solving task. Participants memorized several instances of a pattern and then solved a sequence extrapolation problem that required the detection and generation of novel instances of that same pattern. We observed large transfer effects from memorization to problem solving for extrapolation problems instantiated within the same symbol domain (near transfer) but not across symbol domains (far transfer). This indicates that the knowledge representation acquired from the memorization training was of *intermediate abstraction*, such that knowledge is abstract in regards to the individual objects in the pattern but domain-specific in regards to the relations between objects. Results are explained via the notion of intermediate abstraction in conjunction with two transfer mechanisms of relational bias and reconstructive recall; the latter is novel.

Keywords: Abstraction, cognitive skill, generative knowledge, learning, memorization, problem solving, sequence extrapolation, sequence learning, transfer

The question how and to what extent knowledge acquired in one context or situation is or can be applied in another context is traditionally discussed in terms of *transfer* (Bransford & Schwartz, 1999; Detterman & Sternberg, 1993; Cormier & Hagman, 1987; Perkins & Salomon, 1989; Royer, 1979; Salomon & Perkins, 1989). Knowledge is the main guide for action, so transfer is the key

---

[1] The preparation of this manuscript was supported by Grant No. N00014-99-1-0929 from the Office of Naval Research (ONR) to the second author. No endorsement should be inferred.

[2] Address for correspondence: Timothy J. Nokes, University of Illinois at Chicago, Department of Psychology, M/C 285, 1007 West Harrison Street, Chicago, IL 60607. Phone: (312) 996-6643, fax: (312) 413-4122. Email: tnokes@uic.edu

to the relation between prior learning and subsequent problem solving. Several authors have pointed out that the static metaphor invited by the term "transfer" is misleading (e.g., Carraher & Schliemann, 2002). We agree, but the field has not yet settled on an alternative term. Throughout this article, we use the term "transfer" in a theoretically neutral way to refer to the fact that prior knowledge is brought to bear on future problems, without implication about the nature of the underlying cognitive processes.

A transfer scenario consists of multiple components. First, there is an *acquisition context* in which some *target knowledge* is acquired via some *training task*. Second, there is some *application context* in which that knowledge is brought to bear on some unfamiliar problem or task, the *transfer task*, via one or more *transfer processes*. Acquisition precedes application in time, so the target knowledge appears as *prior knowledge* when viewed from the perspective of the application context. Each of these components can vary along multiple dimensions. For example, the acquisition context can vary in how much and what kind of training it provides, and the transfer task can vary in how similar it is to the training task.

Empirical research has revealed something about the determinants of transfer. Bransford and Schwartz (1999) list the following variables as boosting transfer: a high degree of initial learning (e.g., Klahr & Carver, 1988; Lee & Pennington, 1993), effortless knowledge retrieval (e.g., Beck & McKeown, 1983), deep conceptual understanding (e.g., Brown & Kane, 1988; Chi, Slotta, & deLeeuw, 1994), learning through problem solving (e.g., Adams, Kasserman, Yearwood, Perfetto, Bransford, & Franks, 1988; Lockhart, Lamon, & Gick, 1988), presenting concepts in multiple contexts (e.g., Gick & Holyoak, 1983), representing problems and solutions at the appropriate level of abstraction (e.g., Chen & Daehler, 1989; Novick, 1988), and using metacognitive strategies (e.g., CTGV, 1994; Shoenfeld, 1985).

Although such findings represent progress in understanding transfer, several review articles point out that transfer studies exhibit a mixture of positive and negative results (Bransford & Shwartz, 1999; Gick & Holyoak, 1987; Salomon & Perkins, 1989; Perkins & Salomon, 1989). While some studies have failed to find large transfer effects where we intuitively expect them (Bassok & Holyoak, 1989; Bransford, Sherwood, Vye, & Reiser 1986; Reed, Dempster, & Ettinger, 1985), others have found transfer effects under particular types of study and test conditions (Brown & Kane, 1988; Carney & Levin, 2000; Rittle-Johnson & Alabali, 1999). The psychological reality of transfer is not in doubt, but there are open questions regarding its nature and determinants.

The complexity of the empirical results suggests that transfer is a heterogeneous phenomenon. Greater clarity might result if we assume that different transfer processes are triggered in different types of transfer scenarios. Consistent with this view, researchers have proposed a variety of cognitive processes that could be involved in transfer.

The variety of transfer processes

The oldest hypothesis about the mechanism of transfer, inherited from philosophy and logic, is that knowledge transfers when it is *abstract*, i.e., it leaves out information about specific features and instead encodes relational structure (Ohlsson, 1993). This idea implies that there are two distinct processes contributing to transfer: *abstracting*, the process by which abstractions are formed, and applying or instantiating – we prefer the term *articulating* (Ohlsson & Lehtinen, 1997) – a previously learned knowledge structure vis-à-vis a new situation or problem. Strong claims for transfer of abstract ideas and principles were advanced early in the 20th century by Judd (1908), Katona (1940), and Wertheimer (1945), among others, but the reports of their studies lack many details about method and data analysis and their results have not been replicated. However, several contemporary studies have also claimed strong transfer effects of abstract concepts and principles (e.g., Brown & Kane, 1988; Carney & Levin, 2000; Chi, Feltovich, & Glaser, 1981; Chi, Slotta, & deLeeuw, 1994; Cheng, 2002; Goldstone & Sakamoto, in press; Ohlsson & Regan, 2001; Rittle-Johnson & Alabali, 1999). In many of these studies, the relevant abstractions were explicitly and deliberately taught to the participants (Carney & Levin, 2000; Cheng, 2002; Rittle-Johnson & Alabali, 1999), as they would have been in an educational context (National Research Council, 1999; Perkins & Salomon, 1989; Schwartz, Lin, Brophy, & Bransford, 1999). Although the term "abstraction" is not as popular as it once was, its close modern equivalent is "schema". Schemas are generally thought of as complex knowledge structures that ignore accidental detail but encode the principled features of their references (Gick & Holyoak, 1980, 1983; Ohlsson, 1993; Marshall, 1995; Rumelhart, 1975; Schank, 1986; Thorndyke, 1984) and that hence can transfer to novel situations that share the same principled features.

Strong claims about the transfer of abstract knowledge have been put forward with respect to implicit learning paradigms (Stadler & Frensch, 1998). For example, studies on *artificial grammar learning* (Reber, 1967, 1989, 1993) provide evidence for unintentional abstraction. A large body of evidence (Altmann, Dienes, & Goode, 1995; Berry, 1997; Manza & Reber, 1997; Reber, 1993; Stadler & Frensch, 1998) shows that people perform better than chance when recognizing novel strings generated by the same grammar as a set of memorized training strings, indicating that they have acquired knowledge of the underlying grammar even though they were not instructed to do so. The mechanism assumed to be operating here is unintentional *induction* of the underlying grammar from the training strings.

There are also transfer mechanisms that do not rely on abstraction. The principle of *identical elements* was originally proposed by Thorndike and Woodworth (1901). Gagne (1962, 1965) later proposed a very similar principle of *vertical transfer*. The claim is that task knowledge is highly modular and that it transfers from task X to task Y precisely to the extent that knowledge units required to solve task X are also required to solve task Y. Singley and Anderson (1989) have reformulated this principle as an identical rules principle, where the 'elements' are interpreted as individual production rules. They and Kieras and Bovair (1986) have shown that in some transfer situations, the amount of procedural-to-procedural transfer can be accurately predicted by counting the number of overlapping production rules. In this mechanism, transfer happens automatically as a side effect of the encoding of the procedural knowledge, in combination with the workings of a rule-based cognitive architecture.

Although abstraction and identical elements are the classical transfer theories, modern cognitive research has proposed several mechanisms that differ significantly from both. The most prominent example is *analogical transfer* (Gentner, Holyoak, & Kokinov, 2001). This process has three parts: retrieving a prior knowledge structure, creating a mapping between it and the current problem or situation, and then using that mapping to create new knowledge structures relevant to the application context. A bewildering variety of mechanisms have been proposed for the mapping process (Carbonell, 1986; Forbus, Gentner, & Law, 1995; Holyoak, Novick, & Melz, 1994; Keane, Ledgeway, & Duff, 1994; Holyoak & Thagard, 1989; Hummel & Holyoak, 1997; Salvucci & Anderson, 2002). Some assume that the mapping is a set of binary relations between the two analogues, others that the mapping creates an abstract knowledge structure or schema, different from either of the two mapped structures, that encodes their shared structure. The empirical evidence for analogical mapping is extensive (Catrambone & Holyoak, 1989; Gentner & Gentner, 1983; Gentner & Toupin, 1986; Gentner, Ratterman, & Forbus, 1993; Gick & Holyoak, 1980, 1983; Keane, 1987; Novick, 1988), but the evidence also shows that even though people are capable of mapping deep relational structures, the retrieval of a useful analogue is nevertheless heavily influenced by matches between surface features of the current problem and past problem solving experiences (Catrambone, 2002; Holyoak & Koh, 1987; Ross & Kilbane, 1997). That is, analogy is perhaps a better explanation for near transfer than for far transfer.

The *knowledge compilation* mechanism proposed by John Anderson and co-workers (Anderson, 1983; Neves & Anderson, 1981) was specifically proposed to explain how declarative knowledge can be brought to bear on problem solving in the context of the ACT-R theory. This computational mechanism includes a deliberate and explicit, step-by-step interpretation of the declarative statement, which generates new production rules as a side effect; those rules are then optimized via rule composition, and the result is a procedural representation of the content of the declarative knowledge, given a particular goal. Empirical support for the knowledge compilation mechanism is modest (Neves & Anderson, 1981). It is noteworthy that the knowledge compilation mechanism has not been retained in later versions of the ACT-R theory (Anderson, 1993; Anderson & Lebiere, 1998).

In past work, we have proposed an alternative approach to declarative-to-procedural transfer via error correction (Ohlsson, 1996). Ohlsson and co-workers (Ohlsson, 1996; Ohlsson, Ernst, & Rees, 1992; Ohlsson & Rees, 1991) have proposed that the role of declarative knowledge is primarily to help the learner identify and correct his or her own errors. The *constraint violation* theory envisions the declarative and procedural components as operating in parallel, and the function of declarative knowledge is to constrain possible problem solutions. When incomplete or faulty procedural knowledge generates undesirable outcomes, these are recognized as violations of those constraints and the responsible rules are revised accordingly. The power of declarative knowledge lies primarily in its ability to help pinpoint the cause of an error, and transfer is the process by which errors are identified and remedied. The constraint violation theory has been shown to generate power law learning curves (Ohlsson, 1996) and to support the design of successful tutoring systems (Mitrovic & Ohlsson, 1999).

The key problem in transfer is the relation between the prior knowledge and the target context. That relation is a function of exactly what is learned in the acquisition context and how that knowledge is represented. Whatever the transfer process, one might argue that the possibility of transfer is a function of how the target knowledge was represented in the first place. This has an obvious implication: If we encourage learners to represent the target knowledge in the same way in which it might be used in future problem solving, this should increase the probability of transfer. This principle of *transfer-appropriate processing* has received support in studies with educationally relevant materials. Morris, Bransford, and Franks (1977) have shown that information gained during study is not equally applicable to all possible subsequent tasks, but applies most effectively to transfer tasks that are congruent with the prior training task. For example, Adams et al. (1988) have shown that people can apply previously acquired declarative knowledge to solve simple insight problems, but only when the prior knowledge was acquired under similar processing conditions as the application situation. Specifically, the participants in the "problem-oriented" acquisition condition were able to transfer their knowledge to insight problem solving whereas participants given the exact same information but in the "fact-oriented" acquisition condition did not.

To summarize, the two classical and opposing theories of abstraction and identical elements have both received conceptual face lifts after the cognitive revolution, the first morphing into schema theory and extended into the domain of implicit learning, the second morphing into the identical rules theory. In addition, cognitive researchers have proposed multiple alternative transfer processes, including analogical mapping, knowledge compilation, and constraint violation. Finally, the principle of transfer-appropriate processing highlights the idea that the operation of any transfer process is strongly affected by how the prior knowledge was encoded during training, with some representations being more appropriate for some transfer tasks than others.

These transfer processes are not mutually exclusive. All of them have at least some support in empirical data. It is plausible that the human brain is capable of carrying out these processes, or processes that are functionally equivalent. Greater clarity with respect to transfer requires that we consider which of these processes might be operating under which conditions. This, in turn, requires that we consider how transfer scenarios differ from each other. We review five dimensions of transfer scenarios that we claim have theoretical relevance. We then situate the empirical study reported in this article with respect to those dimensions.

Dimension of transfer scenarios

Transfer within and across domains. It has always seemed intuitively obvious that tasks can be ordered according to the distance from a given task, and that transferring knowledge from a training task A to some transfer task B becomes more difficult when the distance between A and B increases. However, finding a workable and psychologically relevant metric for the degree of distance between two tasks has proved difficult. In fact, the philosopher Nelson Goodman has pointed out that a general problem associated with determining the *overall similarity* between any two objects or situations is that those objects or situations may differ from one another in one of any number of respects (Goodman, 1972; see Medin, Goldstone, & Gentner, 1993 for the psychologists reply).

In the absence of a quantitative metric, it is useful to differentiate between transfer *within a domain* of knowledge, defined in terms of its typical objects, properties, and relations, and transfer *across domains.* Transfer across domains involves relating problems that do not share types of objects or predicates, as in transferring from computer science to cognitive psychology. The natural expectation is that transfer across domains is more difficult. For purposes of this article, we refer to within-domain transfer as *near transfer* and cross-domain transfer as *far transfer,* without assuming that this distinction can be developed into a meaningful quantitative metric. An example of near transfer is the application of previously acquired algebra knowledge to solve various algebra word problems (Clement, 1982; Reed, 1987, 1993) and the transfer of text-editing procedures to learning new text-editing programs (Singley & Anderson, 1989). In contrast, an example of far transfer is the application of knowledge acquired in the domain of mathematics to support problem solving in the domain of physics (Bassok, 1990; Bassok & Holyoak, 1989). The identical elements principle is primarily a process for transfer within domains, because it is unlikely that tasks in two separate domains share any identical elements. Cross-domain transfer via abstraction requires a higher level of abstraction than near transfer, which is presumably more difficult to attain. Analogical mapping is, in principle, equally applicable to transfer across as within domains, as long as an appropriate cross-domain source can be retrieved from memory. The effect of transfer appropriate processing is to decrease the transfer distance.

Declarative versus procedural knowledge. Another dimension along which transfer scenarios vary is whether the knowledge to be transferred is declarative or procedural, and also whether that knowledge, once transferred, is to act as declarative or procedural. Most research on transfer has focused on the procedural-to-procedural case, where some procedural knowledge learned during the acquisition phase is adapted to a transfer task. The identical rules theory is the prototypical theory of this sort (Anderson, Greeno, Kline, & Beasley, 1981; Kieras & Bovair, 1986; Singley & Anderson, 1989). However, knowledge compilation and constraint violation operate primarily on declarative knowledge to bring it to bear on a future problem, i.e., they focus on the declarative-to-procedural case (Neves & Anderson, 1981; Mitrovic & Ohlsson, 1999). Analogical mappings can be constructed between either declarative or procedural knowledge structures. For example, in research on the so-called convergence problem, it is often implicitly assumed that the mapping holds between the two problem statements, an example of declarative-to-declarative transfer (Gick & Holyoak, 1980; 1983; Holyoak & Koh, 1987). In a different application, analogical mappings can be seen as holding between problem solving strategies (e.g., rule sets or individual rules; Anderson & Thompson, 1989) or between solution paths (Carbonell, 1986), both examples of procedural-to-procedural transfer. Finally, the procedural-to-declarative case is primarily discussed in the literature on cognitive development (e.g., Karmiloff-Smith, 1992). The principle of transfer-appropriate processing applies to all of these cases.

Explicit knowledge representation. Transfer processes differ whether they presuppose that the learner has access to an explicit representation of the target knowledge or not. The knowledge compilation mechanism proposed by John Anderson and co-workers (Anderson, 1983; Neves & Anderson, 1981) is the prototypical example of a transfer process that requires an explicit representation of the relevant knowledge. Analogical mapping is usually thought to require recall of a prior analogue. Explicit representations are of course typically available in educational and training contexts. However, many everyday situations do not carry their lessons on their sleeves, as it were, but the latter have to be constructed by the learner. Induction of commonalities is one process that has been proposed to operate in such implicit learning situations, and it implies that transfer from implicit learning scenarios is mediated by abstractions. Sharing of identical production rules is also an implicit process.

Advance knowledge of the transfer task. The principle of transfer-appropriate processing presupposes that the learner knows something about the application context while still in the acquisition context (or, alternatively, that a benevolent teacher or coach deliberately constructs the acquisition context in such a way that the learner unwittingly processes the learning materials in a transfer-appropriate way). Most other transfer principles implicitly assume that the prior knowledge is encoded in whatever format is natural in the acquisition context, without advance knowledge of the application context. This is certainly true of many everyday learning situations.

Generativity of target performance. Transfer scenarios differ with respect to the nature of the transfer task. Almost all the transfer processes discussed above presuppose that the target performance is to solve a problem or to perform some relatively complicated, multi-step task such as word processing. That is, they assume that the purpose of the target knowledge is to generate complex behavior.

In contrast, the methodology used in most studies of implicit learning and artificial grammar learning do not require the participants to generate an organized, temporally extended sequence of actions. Instead, they typically assess what has been learned with various judgment tasks, most typically recognition tasks. In the typical artificial grammar procedure, the knowledge gained during the training phase is assessed by binary judgment tasks in which the participants decide whether novel symbol sequences are of the same type as the previously memorized sequences. This test procedure is very sensitive. However, performance on such single-response judgment tasks do not tell us much about the cognitive function of the knowledge acquired in implicit learning paradigms. For example, to what extent can the participants in artificial grammar learning experiments generate sequences that conform to the grammar they learned? The same question applies to work in the incidental learning (Postman, 1964; Wattenmaker, 1999) and implicit memory paradigms (Richardson-Klavehn & Bjork, 1988; Schacter, 1987). We do not know to what extent the abstract knowledge acquired in those paradigms is generative, i.e., whether it can support problem solving, planning, and other higher-order cognitive processes.

Summary. Transfer scenarios differ with respect to a variety of dimensions, including the transfer distance, the type of knowledge involved, whether the prior knowledge is explicit, whether the learner has advance knowledge of the application context, and the degree of generativity required by the transfer task. These dimensions create a five-dimensional matrix of transfer scenarios, each cell corresponding to some set of values on these five dimensions. Each type of scenario triggers a different subset of processes from the learner's repertoire of transfer processes. This perspective suggests that greater clarity and systematicity in transfer research can be achieved if researchers grouped empirical results according to the properties of the transfer scenarios in which the results were obtained. Rather than performing studies that pit different transfer theories against each other, we need to investigate the transfer behavior that occurs in particular cells in the

matrix. In addition, we need to focus less on the existence of transfer, and more on questions about transfer distance and the amount of transfer. If transfer effects turn out to be smaller or larger in some type of scenario than in another, this fact might give us important clues as to the strength and prevalence of the various hypothesized transfer processes. Unfortunately, many transfer studies do not report effect sizes (e.g., Brown & Kane, 1988; Carraher & Schliemann, 2002; Gick & Holyoak, 1983; Rittle-Johnson & Alabali, 1999).

The present study

Interest in transfer is, in part, driven by the persistent intuition that declarative knowledge does indeed facilitate action vis-à-vis unfamiliar problems or situations. Everyday life and common educational practices (e.g., the organization of textbook chapters) seem inexplicable unless we believe that declarative knowledge learned in one context routinely enter into and influence the processing of future problems and situations. However, the available empirical data base regarding declarative transfer is currently too small to serve as a basis for theorizing. The nature of declarative transfer has been an issue for some time in the area of mathematics learning (Hiebert, 1986), but Rittle-Johnson and Alibali (1999) nevertheless recently summarized the field by saying that Perry (1991) was "the only study that showed declarative transfer in mathematics" (p. 176). In many transfer studies, it is often unclear whether the prior knowledge on which transfer processes might operate was declarative or procedural to begin with. This is often the case with studies on analogical transfer because the target knowledge is associated with solving a problem, so there are both declarative representations of the problem and the procedural attachments for how to solve the problem (Chen, 2002). Basic information about transfer effects in situations which clearly involve declarative transfer would be useful.

We choose to study declarative transfer in a situation in which the target knowledge is not explicitly presented to the learner. Although educational contexts are characterized by the fact that declarative knowledge is typically presented in codified form (e.g., as explicitly stated facts or principles), many, perhaps most, everyday learning situations do not come with their lessons written on their sleeves, as it were. They leave it to the learner to figure out which lesson to draw from the relevant activity or experience. If transfer of declarative knowledge is as ubiquitous as intuition holds, then even incidentally acquired declarative knowledge must transfer. However, we know little about the question of transfer distance with respect to incidentally acquired declarative knowledge. Does such knowledge transfer across domains? If it does transfer, how large is the transfer effect?

Furthermore, we choose to address the issue of declarative transfer in the context of situations in which the learner does not know what the transfer task looks like, or indeed that there is a transfer task. They acquire the relevant declarative knowledge in a training scenario that does not refer in any way to the intended transfer task. Many of the studies of declarative transfer have been carried out in educational contexts in which the learners know for what kind of future tasks the knowledge will be useful. In these kinds of situations, deliberate preparation for transfer is possible. We know less about declarative transfer in situations that do not allow deliberate transfer-appropriate processing. Yet, if the application of prior knowledge to future problems and situations is ubiquitous, as intuition would have us believe, then transfer without deliberate transfer-appropriate processing must be the rule. We seldom know exactly which types of tasks we will encounter in the future, and in most situations, we acquire declarative knowledge as a side effect of activity or in a local context that is framed in terms of acquiring the target knowledge, rather than in terms of preparing for some specified future task.

Finally, we are interested primarily in the transfer of incidentally acquired declarative knowledge to problem solving situations. Research on implicit learning, especially artificial grammar learning, strongly support the hypothesis that people extract implicit, abstract knowledge about such grammars by memorizing strings generated by those grammars, but we do not know to what extent the knowledge acquired in this way transfers to problem solving. As we reviewed above, most implicit learning studies use recognition judgments as transfer tasks. We are interested in transfer tasks that require the learners to generate a series of coordinated actions or responses.

In summary, our aim was to investigate a transfer scenario in which there is some well-defined declarative knowledge that the participants in our study would not have had before they participated. The target knowledge should be clearly differentiated from the procedural knowledge needed to solve the transfer task. The function of the declarative knowledge in solving the transfer task should be clearly specifiable. Furthermore, there should be some reliable technique by which we can teach the target knowledge to the experimental participants implicitly, without an explicit codification or presentation. We choose not to inform the participants about the nature or existence of the transfer task thus preventing transfer-appropriate processing. Finally, the transfer task should be generative in character.

We ask four questions about this type of transfer scenario: First, does within-domain declarative transfer occur? That is, are there measurable effects on problem solving performance? Second, does this kind of transfer scenario support cross-domain transfer? Third, are there significant individual differences with respect to how people perform in this transfer scenario? Fourth, how large are the transfer effects? Are they large enough to make a difference in everyday situations?

To answer these questions, we adapted the letter sequence extrapolation task introduced by Thurstone and Thurstone (1941) as our target task. In the traditional form of the task, the problem solver is shown a patterned sequence of letters, such as

M A B M C D

He or she is then asked which letter comes next in the sequence. Letter sequence extrapolation tasks have been studied from a cognitive point of view by Greeno and Simon (1974), Klahr and Wallace (1970), Restle (1970), Restle and Brown (1970), Simon (1972), Simon and Kotovsky (1963), and others. To strengthen the assessment of generativity, we modified the task by asking our participants to continue the sequence for an entire period of the underlying pattern instead of merely supplying a single next letter. The solution to the above example problem would thus be

M E F

This problem is quite simple, but sequence extrapolation problems can be made arbitrarily complex. The reader might want to try to extrapolate the sequence

A C Z D B Y Y D F X G E W W

to eight places. Performance on this type of task depends on a well defined part of the participants' prior knowledge: the alphabet.

A solution to a sequence extrapolation problem requires two main cognitive processes. First, the problem solver must analyze the given portion of the sequence to discover the embedded pattern. It is not sufficient to memorize the given part of the sequence. Being able to recite "M A B M C D" does not in and of itself provide the ability to generate the continuation "M E F". Successful performance requires that the problem solver detects and encodes the relations between the positions in the given sequence. Second, the problem solver must figure out which letters constitute the correct continuation. Even with knowledge of the relevant relations, the required inferences are not trivial. They can be conceptualized as multiple miniature analogies: If position P1 is analogous to position P3, P2 is analogous to P4, the letter in position P1 has relation R to the letter in P2, and if the letter in position P3 is X, then the letter in position P4 should have relation R to X. Each position in the extrapolated sequence requires an inference of this general sort.

Sequence extrapolation tasks are obviously easier to solve if one already knows the pattern embedded in the given part of the sequence. We can therefore study declarative-to-procedural transfer by providing participants with the opportunity to acquire the relevant pattern in a prior study phase, and then observe the effect on their performance on a subsequent extrapolation task. In the study reported here, the learning phase consisted in committing to memory several symbol sequences that embodied the same pattern as the target extrapolation task. Committing patterned sequences to memory is the most commonly used task in the implicit learning paradigms that provide the bulk of the evidence for abstract representations of underlying patterns.

We used sequences of double digit numbers rather than letters as our study material. There are four times as many double digit numbers as there are letters, so we could construct study sequences that did not overlap with (i.e., did not share any symbols with) either each other or with the transfer problems. In this situation, facilitation of performance on the target problem cannot be mediated by recurring surface features (Perruchet & Gallego, 1997; Servan-Schreiber & Anderson, 1990), because there are no recurring surface features.

We used two types of target problems. The *near transfer problems* were expressed in terms of the same type of symbol – double digit numbers – as the study material. Like letter sequence extrapolation problems, number sequence extrapolation problems depend on a well defined part of the participants' prior knowledge: place-value notion for the natural numbers. The *far transfer problems* were expressed in terms of letters. If the knowledge gained during the study phase is tied to the training sequences themselves, we should see no facilitation of performance on either type of transfer task, as compared to a control group that studies random sequences. If the knowledge gained during study does transfer within the domain, i.e., to the near transfer problem, the question remains whether it also transfers to the far transfer tasks.

In short, the purpose of the study reported here was not to determine *whether* people can acquire information about an embedded pattern by studying the training sequences. Instead we asked, *given* that relevant knowledge was acquired from study of the training strings, what are the characteristics of that knowledge? Does it transfer to deliberate problem solving? If so, how far does it transfer? Is it limited to the same domain, or does it transfer across a domain boundary? Also, how large is the transfer effect? Is it large enough to be important in a real situation? Although we employ a classical treatment-control experimental paradigm, our intent is more descriptive than hypothesis testing.

Method

Table 1. Two sequence extrapolation problems expressed in both letters and numbers.

Problem	Given letter/number sequence	Correct 8-step extrapolation
1a	B D X E C Z E G X H F Z	H J X K I Z K M
1b	25 27 47 28 26 49 28 30 47 31 29 49	31 33 47 34 32 49 34 36
2a	C D B E A M D E C F B N	E F D G C O F G
2b	63 64 62 65 61 73 64 65 63 66 62 74	65 66 64 67 63 75 66 67

Participants. Ninety-eight undergraduate students from the University of Illinois at Chicago participated in return for course credit.

Materials. The transfer tasks were sequence extrapolation problems with a periodicity of six items. To enable the participants to detect the pattern, the given segments were 12 items long. That is, they covered two complete iterations of the pattern. Each task was instantiated in both numbers (near transfer) and letters (far transfer); see Table 1. Both the number and letter problems follow the exact same pattern for each transfer task. The patterns in these extrapolation problems were invented specifically for this experiment but are similar to those used by Kotovsky and Simon (1973).

There were 24 training strings consisting of 12 double digit numbers, twelve strings for each problem. The twelve strings associated with a problem followed the same pattern as the given sequence for that problem; see the Appendix for examples. In addition, there were 24 strings of random double digit numbers used in the control (random training) condition.

Each participant received a test booklet with two parts. Within each part, there were twelve sheets presenting the training strings, twelve blank recall sheets, one sheet presenting the sequence extrapolation problem, and one blank sheet to assess the participants' knowledge of the pattern. The order of the two parts was counter-balanced across conditions.

Design. The participants were randomly assigned to one of four groups created by pairing patterned training versus random training with near transfer versus far transfer problems: *patterned near* (n = 26), *patterned far* (n = 27), *random near* (n = 21), or *random far* (n = 24).

In the patterned training groups, the participants memorized the strings that conformed to the same patterns as those in the transfer problems; see Appendix A. In the random training groups, the participants memorized the random number sequences.

In the near transfer groups, the transfer problems were number sequence extrapolation problems; see Table 1, problems 1b and 2b. That is, the problems were expressed in the same type of symbols as the training strings. In the far transfer groups, the transfer problems were letter sequence extrapolation problems; see Table 1, problems 1a and 2a. That is, the problems were expressed in a different type of symbol than the training strings.

Procedure. The participants were tested in groups of 25. The procedure consisted of two *cycles.* Each cycle consisted of training followed by problem solving. The participants memorized and recalled twelve number strings, one by one. They were given 60 seconds to memorize each string. They were then told to turn the page and write down the string. This procedure was repeated through the twelve training strings. Next, the participants were told to turn the page and solve the sequence extrapolation problem. In the two patterned training groups the participants were told that the memorization strings may help them solve the problem. They were given 5 minutes to solve the problem. They were then asked to turn the page and describe the pattern in the problem sequence as best they could. The second cycle proceeded in the same way. The procedure took approximately 70 minutes.

Results

The results are divided into four sections. In the first section we examine the participants' performance on the string memorization task. The purpose of the analyses in that section is to determine whether the patterned training groups learned anything about the pattern during training. We are interested in what effect knowledge of the pattern might have on subsequent problem solving, so to verify that the participants did indeed acquire some knowledge of the pattern is an important manipulation check. In the second section we assess the effect of training on problem solving, presenting both mean problem solving performance as well as position-by-position analyses to determine how far the acquired knowledge transferred. In the third section we address the problem that different participants might have acquired more or less knowledge of the pattern during training. We select those participants that performed best during training and make median-split comparisons between those with high and low training performance. These results turn out to confirm the outcomes of the main analyses in section two. Finally, in the fourth section we turn to the question of the size of the observed effects. We present multiple measures of the amount of transfer obtained from the memorization task to the problem solving task.

Section 1: Memory Performance. The first question is whether participants in the patterned training condition learned the pattern embedded in the training strings. Knowledge of the pattern can be used to learn and reconstruct the number sequences so it should improve recall performance. Hence, the participants in the patterned training condition should perform better on the memorization task than the participants in the random training condition.

The *memory score* for each participant was the number of double digit numbers correctly recalled for each training trial. Because there were 12 numbers to memorize per string, the memory score varied between 0 and 12. Figure 1 shows the mean memory scores on trials 1 – 12 for both patterned and random training groups. Alpha was set to .05 for all subsequent tests and effect sizes (eta squared, η^2) were calculated for all main effects, interactions, and main comparisons. Cohen (1988) has suggested that effects be regarded as small when $\eta^2 < .06$, as medium when $.06 < \eta^2 < .14$, and as large when $\eta^2 > .14$.

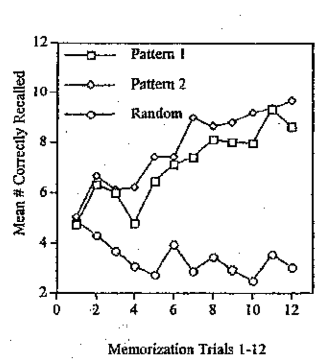

Figure 1. Mean memory scores on trials 1-12 for patterned and random training.

A one-way (training, patterned vs. random) between subjects analysis of variance (ANOVA) revealed an effect for training. The patterned training group performed significantly better than the random training group, F (1, 96) = 88.28, MSE = 4.33, $p < .05$, indicating that the former benefited from the patterns embedded in the training sequences. The size of the effect was quite large, $\eta^2 = .48$. As Figure 1 shows, this effect is present for both pattern types used in the patterned training but not for the random training group. In fact, the latter decline slightly, presumably due to memory interference.

We conducted two additional analyses. To further document that the participants learned during training, we conducted a linear trend ANOVA across trials 1 – 12 for each training group. The patterned training group showed significant positive linear trends across trials for both patterns 1 and 2, F (1, 52) = 34.01, MSE = 25.33, $p < .05$, $\eta^2 = .40$, and F (1, 52) = 70.68, MSE = 18.03, $p < .05$, $\eta^2 = .58$ respectively. In contrast, the random training group showed a significant negative linear trend across trials, F (1, 44) = 23.06, MSE = 4.21, $p < .05$, $\eta^2 = .34$. Taken together, this and the previous analysis provide strong evidence that the patterned training groups learned something about the pattern during training.

We also conducted a one-way repeated measures ANOVA to investigate the effect of pattern type (1 vs. 2) on memory performance for the patterned training group. There was an effect of pattern, F (1, 52) = 7.06, MSE = 2.20, $p < .05$, indicating that strings that instantiate pattern 2 were easier to learn than strings that instantiate pattern 1. This effect was of medium size, $\eta^2 = .12$.

Finally, a t-test was conducted to compare the training performance of participants assigned to the near transfer group to those assigned to the far transfer group. The two groups did not significantly differ in training performance, t (96) = 1.59, *ns*, which makes their problem solving performance comparable.

Section 2: Transfer to Problem-Solving. The second question is whether the participants in the patterned training groups performed better on the problem solving tasks. The *problem solving score* was the number of letters or numbers correctly extrapolated in each problem solving task. Because the participants were asked to extrapolate the sequence to eight places, their problem solving scores varied from 0 to 8.

A 2 (training, patterned vs. random) by 2 (transfer, near vs. far) by 2 (problem, 1 vs. 2) mixed ANOVA revealed no effect for training, $F\ (1, 94) = 1.03$, $MSE = 15.33$, *ns*, $\eta^2 = .01$. However, there was a medium-sized main effect of transfer, $F\ (1, 94) = 10.42$, $MSE = 15.33$, $p < .05$, $\eta^2 = .10$, indicating that participants in the near transfer groups performed better than participants in the far transfer groups. There was also a large main effect of problem, $F\ (1, 94) = 16.99$, $MSE = 6.34$, $p < .05$, $\eta^2 = .15$, indicating that problem 2 was easier than problem 1.

In addition, the interaction of training by transfer was marginally significant, $F\ (1, 94) = 3.94$, $MSE = 15.33$, $p = .05$, $\eta^2 = .04$, indicating that the advantage for participants in the patterned group was larger on near transfer problems than on far transfer problems. Figure 2 (panel a) shows the mean problem solving scores for the patterned and random groups on the near and far transfer problems. Main comparisons show that the patterned group performed significantly better than the random group on the near transfer problems but not on the far transfer problems, $F\ (1, 94) = 4.31$, $p < .05$, $\eta^2 = .04$, and $F\ (1, 94) = .49$, *ns*, $\eta^2 = .005$ respectively. These results indicate that participants in the patterned group benefited from training only when solving the near transfer problems.

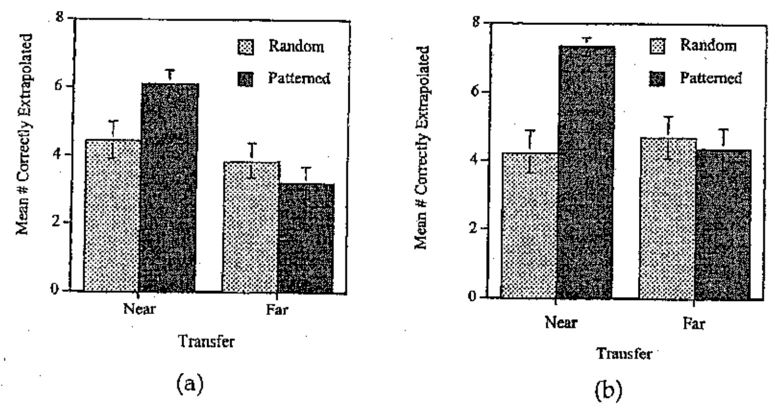

Figure 2. Panel (a) shows mean problem solving scores for all participants in the patterned and random groups on near and far transfer problems. Panel(b) shows mean problem solving scores for top 2/3 of training participants on near and far transfer problems.

Table 2. Number of participants to answer each position correctly for both near transfer and far transfer problems.

Training	1	2	3	4	5	6	7	8
Problem 1 — # of subjects to solve correctly in problem solving, Position								
Patterned								
Near ($n = 26$)	17	15	22	17	14	22	15	14
Far ($n = 27$)	7	7	13	8	6	12	7	6
Effect	*	*	*	*	*	*	*	*
Random								
Near ($n = 21$)	9	9	14	9	8	13	10	8
Far ($n = 24$)	8	9	15	6	5	12	8	7
Effect	-	-	-	-	-	-	-	-
Problem 2 — # of subjects to solve correctly in problem solving, Position								
Patterned								
Near ($n = 26$)	23	23	23	23	23	24	21	20
Far ($n = 27$)	13	14	13	13	14	15	12	12
Effect	*	*	*	*	*	*	*	*
Random								
Near ($n = 21$)	15	14	15	14	14	14	11	9
Far ($n = 24$)	15	15	15	14	15	13	14	13
Effect	-	-	-	-	-	-	-	-

Note: * = significant, p < .05; - = non-significant, p > .05

Table 3. Percentage of participants to solve any given problem position correctly as a function of transfer.

	Training Condition	
Transfer	Patterned	Random
Near	76%	55%
Far	40%	48%

Note: Chi-Square, $\chi^2\ (1, N = 219) = 3.33$, $p = .068$.

To further investigate the interaction of training by transfer, we examined the number of participants from each training group to correctly extrapolate each position of the solution for both near and far transfer problems. Table 2 shows the number of participants to correctly solve positions 1-8 of problems 1 and 2 for both patterned and random groups as a function of transfer. Chi square tests revealed that for patterned training, significantly more participants solved near transfer (number) problems than far transfer (letter) problems. This was true for all eight positions in both problems 1 and 2. In contrast, random training participants did not significantly differ on any one position when solving near versus far transfer problems.

In sum, participants in the patterned training group who solved near transfer problems performed better than those solving far transfer problems and those who received random training. Table 3 shows the percentage of participants to solve any given position correctly for both patterned and random groups as a function of transfer.

Section 3: Problem Solving as a Function of Individual Differences in Training. It is possible that our participants differed in their ability to memorize the training strings. For present purposes, we are interested in the effects of the string learning on subsequent problem solving. That is, we are not interested in *whether* string learning was successful, but in what its effects are *when* successful. Hence, participants who failed the string learning task dilute the effect we are interested in. A look at the frequency distribution of the memory scores confirmed that this is a potential problem; see Figure 3 panel a. The distribution was strongly bimodal, with a minimum at 5.9, close to the mid-way point on the memory score scale. Hence, the participants fell into two different groups: those who performed well on the memorization task (approximately 2/3 of the participants) and those who performed poorly (approximately 1/3 of the participants). In contrast, the frequency distribution of the memory scores for the random training group was not bimodal but approximately normal; see Figure 3 panel b.

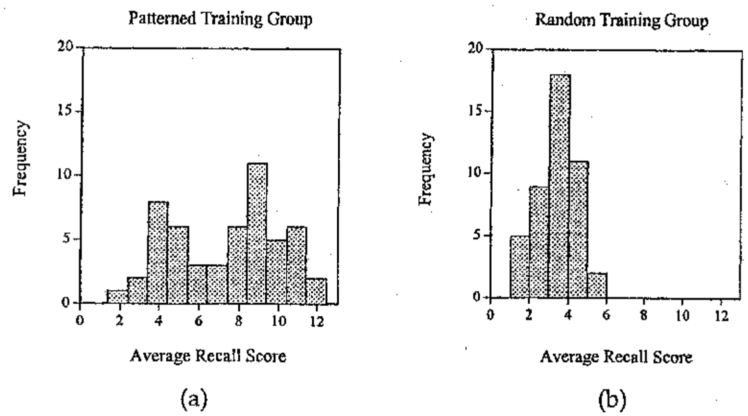

Figure 3. Frequency distributions for mean recall scores for both training groups.

To investigate the effect of memorization on problem solving for participants for whom the pattern learning was successful, the same analyses as before were repeated for the high memory participants. We compared the top 2/3 of patterned training participants to the top 2/3 of random training participants on problem solving performance. A 2 (training, patterned vs. random) by 2 (transfer, near vs. far) by 2 (problem, 1 vs. 2) mixed ANOVA revealed a medium-sized main effect for training, $F(1, 61) = 5.40$, $MSE = 11.52$, $p < .05$, $\eta^2 = .08$, indicating that participants in the patterned training group performed significantly better than the participants in the random training group. There was a medium-sized main effect of transfer, $F(1, 61) = 4.54$, $MSE = 11.52$, $p < .05$, $\eta^2 = .07$, indicating that participants solving near transfer problems performed better than participants solving far transfer problems. There was also medium-sized main effect of problem, $F(1, 61) = 7.15$, $MSE = 6.71$, $p < .05$, $\eta^2 = .10$, indicating that participants performed better on problem 2 than problem 1.

Table 4. High – low memory spilt for each position of the pattern and the corresponding number of participants to correctly extrapolate that position in problem solving.

Problem 1								
Training	# of subjects to solve correctly in problem solving Position							
	1	2	3	4	5	6	7	8
Patterned								
Low (n = 27)	6	3	13	6	1	13	4	1
High (n = 26)	18	19	22	19	19	21	18	19
Effect	*	*	*	*	*	*	*	*
Random								
Low (n = 23)	6	7	13	6	4	10	6	8
High (n = 22)	11	11	16	9	9	15	12	7
Effect	-	-	-	-	-	-	+	-

Problem 2								
Training	# of subjects to solve correctly in problem solving Position							
	1	2	3	4	5	6	7	8
Patterned								
Low (n = 27)	13	14	12	13	14	15	10	9
High (n = 26)	23	23	24	23	23	24	23	23
Effect	*	*	*	*	*	*	*	*
Random								
Low (n = 23)	17	15	15	12	13	12	14	12
High (n = 22)	13	14	15	16	16	15	11	10
Effect	-	-	-	-	-	-	-	-

Note: * = significant, p < .05; + = marginal significance, .06 > p > .05; - = non-significant, p > .05

In addition, there was a medium-sized interaction of training by transfer, $F(1, 61) = 8.16$, $MSE = 11.52$, $p < .05$, $\eta^2 = .12$, indicating that the advantage for the patterned training group on near transfer problems was larger than for far transfer problems. Main comparisons show that the patterned group performed significantly better than the random group on near transfer problems but not on the far transfer problems, $F(1, 61) = 13.60$, $p < .05$, $\eta^2 = .19$, and $F(1, 61) = .14$, *ns*, $\eta^2 = .002$ respectively. Figure 2 (panel b) shows mean problem solving scores for each training group as a function of transfer. These results strongly support those shown in panel a, indicating that when pattern learning was successful large facilitation effects were obtained on near transfer problems. In short, the results for the top two-thirds of participants replicate the results for the entire sample, but with larger effect sizes.

To further investigate the relation between memorization and problem solving, we compared mean memory performance at each position of the sequence to the number of participants who correctly extrapolated that particular position during problem solving. In this analysis we used all participants and they were classified as either high or low memory based upon a median split of memory scores for each training group. Median splits were calculated at each position of the pattern and the number of participants to correctly extrapolate the corresponding position for both high and low memory groups was assessed. Table 4 shows the number of participants to correctly solve positions 1-8 of problems 1 and 2 for both patterned and random groups as a function of memory.

As Table 4 shows, for patterned training the high memory group performed better at each position in the problem solving task than the low memory group for both problems 1 and 2, while no such pattern was present for participants given random training. Chi square tests confirmed that for patterned training, significantly more high memory participants solved corresponding extrapolations than low memory participants for all eight positions of both problems. In contrast, the high memory group given random training only performed (marginally) significantly better than the low memory group on one of eight positions for problem 1 and zero out of eight positions for problem 2.

In addition, we compared high memory participants who received patterned training to high memory participants who received random training for both near and far transfer problems. Each transfer problem (i.e., near vs. far) was analyzed separately which reduced the total number of participants by half since the type of transfer problem was a between subjects variable. Table 5 shows the number of high memory participants to correctly solve positions 1-8 of problems 1 and 2 for both near and far transfer problems as a function of training. Chi square tests show that significantly more high memory participants given patterned training solved the corresponding extrapolations than high memory participants given random training for all positions of problems 1 and 2 but only on the near transfer problems. In

contrast, no difference was found between training groups on far transfer problems.

Table 5. Number of high memory participants to correctly extrapolate each position for both near and far transfer problems as a function of training.

Problem 1	# of subjects to solve correctly in problem solving							
Problem Type	Position							
	1	2	3	4	5	6	7	8
Near Transfer								
Patterned ($n = 13$)	12	12	13	12	12	13	12	12
Random ($n = 11$)	5	5	6	4	4	8	6	4
Effect	*	*	*	*	*	*	*	*
Far Transfer								
Patterned ($n = 13$)	5	5	9	6	5	9	5	5
Random ($n = 12$)	6	6	10	6	5	9	6	4
Effect	-	-	-	-	-	-	-	-

Problem 2	# of subjects to solve correctly in problem solving							
Training	Position							
	1	2	3	4	5	6	7	8
Near Transfer								
Patterned ($n = 13$)	13	13	13	13	13	13	13	13
Random ($n = 11$)	6	7	7	7	7	7	4	4
Effect	*	*	*	*	*	*	*	*
Far Transfer								
Patterned ($n = 13$)	9	9	10	9	9	11	9	9
Random ($n = 12$)	8	8	9	9	10	9	8	7
Effect	-	-	-	-	-	-	-	-

Note: * = significant, $p < .05$; - = non-significant, $p > .05$

Table 6. Percentage of participants to solve any given problem position correctly as function of memory performance.

	Training Condition	
Memory Group	Patterned	Random
Low	34%	46%
High	82%	57%

Note: Chi-Square, $\chi^2(1, N = 219) = 15.22$, $p < .05$

Table 7. Measures of transfer for patterned learning groups on near transfer problems.

Concept	Formula	Value	
		All Participants	Upper 2/3
Percent of transfer improvement normalized by control group	$T\%_{Improvement} = (E - C / C) \times 100$	38%	72.5%
Percent of transfer improvement normalized for practice	$T\%_{Improvement} = (E_{c2} - C_{c2} / C_{c2} - C_{c1}) \times 100$	158%	120%
Percent of transfer improvement of total learning possible	$T\%_{Total\ Learning} = (E - C / {}_{Perform\ limit} - C) \times 100$	47%	83%
Effect Size			
Cycle 1	Standardized Mean Difference	$\delta = .46$	$\delta = 1.44$
Cycle 2	$\delta = (\mu_E - \mu_C) / S_{pooled}$	$\delta = .63$	$\delta = 1.11$

Note: E = patterned training group, C = random training group, $_{c1}$ = cycle 1, $_{c2}$ = cycle 2. Upper 2/3 refers to the top two-thirds of performers on the memorization task.

In sum, participants in the patterned training group classified as high memory performed better than those classified as low memory and those who received random training. Table 6 shows the percentage of participants to solve any given position correctly for both patterned and random groups as a function of memory. These analyses show support for the claim that what was learned during the training phase influenced cognitive processing during problem solving, but only for near transfer problems

Section 4: Amount of Transfer. We now turn to the question of the size of those effects. In order to quantify the amount of transfer observed from the memorization task to the problem solving task we calculated three measures of transfer; see Table 7 for a summary. Our first transfer measure, adapted from Singley and Anderson (1989), quantifies the amount of improvement of the experimental (patterned training) group the over control (random training) group normalized by control group performance, $T_{\%\ Improvement} = (E - C / C) \times 100$. This transfer measure revealed a 38% improvement for the patterned training group compared to control on near transfer problems. We also calculated this measure for the upper 2/3 of training participants and found a 72.5% improvement on near transfer problems.

Our second measure of transfer quantifies the improvement of the experimental group compared to control at cycle 2 normalized by the improvement of the control group from Cycle 1 to Cycle 2 on the transfer task, $T_{\%\ Improvement} = (E_{c2} - C_{c2} / C_{c2} - C_{c1}) \times 100$. This measures the effect of training condition over and beyond the improvement caused by solving the target task twice. This transfer measure revealed a 158% improvement for patterned training group on the near transfer problems for cycle 2. The same measure was calculated for the upper 2/3 of training participants revealing a 120% improvement on near transfer problems.

Our third measure of transfer, also adapted from Singley and Anderson (1989), quantifies the amount of transfer given the total amount of learning possible; $T_{\%\ Total\ Learning} = (E - C / {}_{Performance\ limit} - C) \times 100$. If one knows the maximum amount of performance possible on the transfer task (in this case 8 correct extrapolations) you can calculate the total amount of learning possible (i.e., performance limit minus control performance). This measure revealed a 47% increase of the total amount of improvement possible for patterned training participants on the near transfer problems. The same measure was calculated for the upper 2/3 of training participants revealing an 83% increase of the total amount of improvement on the near transfer problems.

Finally, we calculated effect sizes for the patterned training group on the near transfer problems for cycles 1 and 2. Effect sizes ranged from medium to large, according to Cohen (1988), for estimates of the standardized mean difference (δ) using the pooled standard deviation. Participants given patterned training performed approximately half a standard deviation better than the random training group on near transfer problems for both cycles 1 and 2 ($\delta = .46$ and $\delta = .63$ respectively). When individual differences were taken into account via the upper 2/3 analysis the effect was even stronger, $\delta = 1.44$ on cycle 1 and $\delta = 1.11$ on cycle 2.

Discussion

Our question was not whether the participants in our study could learn the patterns embedded in the training strings by memorizing those strings. Instead, our questions were: If the participants did learn those patterns, could they transfer what they had learned to a problem solving task that required them to generate a novel, coordinated sequence of responses? If so, how far did the knowledge transfer? Could the participants transfer the patterns across a domain boundary? Finally, how large was the transfer effect? Was it large enough to warrant theoretical attention and to be important in a nonlaboratory situation?

Interpretation

Our participants did learn something about the patterns embedded in the training strings. The evidence for this is their higher recall of the sequences, compared to the group that memorized random sequences, and the gradual increase in recall accuracy across the twelve training trials for both patterns (see Figure 1). The latter increase did not appear in the random group, so it was not due to practice in memorizing strings. These findings show that the participants in the patterned training group gradually encoded the recurring relations in the training sequences.

Was the participants' knowledge of the underlying patterns generative? Neither the near transfer nor the far transfer tasks could be solved by reciting or recalling the training strings. The sequence of symbols that constituted the answer to the extrapolation problems was not identical to any of the training strings, but had to be constructed. Could the participants draw upon what they had learned to generate those sequences?

The participants who studied the patterned training strings had higher problem solving scores on the near transfer tasks than those who studied the random training strings (see Figure 2). This effect is accentuated if we focus on those 2/3 of the participants who performed best on the training tasks (see Figure 2 b). Even stronger evidence comes from the position by position analyses. The participants in the patterned training condition whose recall accuracy was above median on the training tasks performed better during problem solving than those whose recall was below the median (see Table 4). This was true for every position in both patterns. No such effect was present in the random training condition.

Neither the given parts of the sequence extrapolation problems nor the solutions to those problems appeared among the training strings. Hence, the fact that the training facilitated the solutions to the near transfer problems implies that the knowledge gained during training was not specific to the particular symbols in the training strings. For the improvement on the near transfer problems to have occurred, the participants must have subjected that knowledge to some type of transfer process.

Did that process also prepare the participants for the far transfer problems? Could they productively apply the acquired knowledge across a domain boundary? Our data suggest that this was not the case. The patterned training groups did not perform significantly better than the random training groups on the far transfer problems (see Figure 2). This conclusion is strongly supported by the position-by-position analyses. The number of participants getting any one position correct during problem solving was consistently larger in the near transfer condition than in the far transfer condition. This was true of both problem 1 and problem 2 (see Table 2). There was no such effect for random training. These effects are accentuated if we focus on those 2/3 of the participants who learned the most in the training phase (see Figure 2b and Table 5). The transfer processes operating in this scenario did not enable the participants to apply the acquired knowledge across the boundary between the domain of numbers and the domain of letters.

Advanced knowledge of the far transfer task did not help in this regard. Every participant went through two cycles of training followed by problem solving (see Method section). In the first cycle, they did not know that they would encounter a sequence extrapolation problem in the course of the experiment. Hence, they had no reason to approach the training task in any other way than as a memory task. However, during the second cycle some subjects might have inferred that there would be yet another extrapolation problem at the end of that cycle, so, they had the opportunity to engage in transfer-appropriate processing. The data suggest that this did not happen. The patterned training groups did improve from cycle 1 to cycle 2. However, so did the random training groups, and by approximately the same amount (see Figure 4). This improvement most likely represents the effect of having solved one prior extrapolation problem, not an effect of changing the strategy for how to commit the training strings to memory.

How much transfer was there? The exact amount depends on how transfer is measured, but on any measure the effect was large. The participants who performed in the upper 2/3 on the training tasks reached 70 – 120% transfer on three different transfer measures. Standard measures of statistical effect size show that the training manipulation moved the score distribution for the patterned training group more than one standard deviation upwards; see Table 6. Unfortunately, there are no widely accepted guidelines for how large a transfer effect must be in order to count as important, either theoretically or practically. However, 70% or more transfer in a problem solving context translates into a substantial reduction of the amount of cognitive processing required to perform the transfer task, and a theory of the acquisition of cognitive skills ought to explain that reduction. Also, moving a score distribution upwards by one standard deviation would count as a substantive improvement in most educational contexts. Hence, we conclude that the transfer processing that was operating in our transfer scenario is of both theoretical and practical significance.

Explanation

By what processes did the knowledge acquired by studying the training strings facilitate performance on the extrapolation problems? A sequence extrapolation problem has two main subtasks: To identify the pattern in the given part of the sequence, and to extrapolate it to the required number of positions. Training might have impacted either or both of these processes.

Research in implicit and artificial grammar learning (Manza & Reber, 1997; Reber, 1993) has demonstrated that the task of memorizing patterned symbol strings generates abstract representations of those patterns, so the abstraction theory provides a natural explanation for how pattern knowledge might have influenced problem solving performance. When the participants studied the training strings, they responded to the recurring relations in string after string. As a consequence, they acquired a schema for the system of relations that made up the underlying pattern. It is highly likely that this prior processing influenced memory access and attention allocation during problem solving. The participants had a higher probability of noticing the crucial relations between the symbols in the given part of the extrapolation problem because they had just studied other symbol strings that contained those relations. As a consequence, they identified the pattern embedded in the given part of the extrapolation problem faster, more accurately and with less cognitive effort, leaving more cognitive capacity and more time to work out the extrapolation inferences.

Because the training strings and the given parts of the extrapolation problems did not share any elements, this effect could only have occurred if the relevant relations were encoded abstractly. If so, why did the participants not improve on the far transfer problems as well? To resolve this puzzle, we distinguish between two levels of abstraction (Ohlsson, 1993). Mental representations at the *object level* of abstraction ignore the individual elements or objects in a relational structure but encode the relations in a domain-specific way. For example, the relation between the two numbers 22 and 23 can be encoded as *next number up, N plus one, add unity,* or in many other, equally number-specific ways. Similarly, the relation between the two elements in the pairs (17, 19), (23, 25) and (34, 36) can be represented as *X plus 2, X increased by 2,* and in other ways that refer to numbers and operations on numbers. A pattern encoded in terms of such number-specific relations is abstract vis-à-vis particular numbers, but it is nevertheless not transferable to a non-numerical domain such as letters without further processing.

For the knowledge to have transferred from the training tasks to the near transfer tasks, it must have been abstract in at least this sense. In addition to learning particular propositions like *the symbol in position n_1 is 15, the symbol in position n_2 is 17,* and so on, the subjects learned propositions like *the relation between the symbol in position n_1 and the symbol in position n_2 is plus 2.* A set of propositions of the latter form is a schema that is abstract in the sense of generalizing over individual symbols but domain-specific in that it is encoded in terms of relational predicates that are unique to the domain of numbers.

A higher level of abstraction is possible because the relational predicates can themselves be conceptualized in domain-independent ways. For example, the relation *next element,* suggested by Simon (1972), applies equally to letters and numbers, so it is more abstract than either *next number* or *next letter* (as long as both *number* and *letter* qualify as a kind of *element*). Similarly, the relation between the two elements in the pairs (17, 19), (23, 25), and (34, 36) can be encoded as *two steps forwards in the relevant sequence, the symbol that follows the one that follows X,* and in other ways that are not unique to numbers but applies equally well to letters.

Evidently, our participants did not encode the pattern at this *relational level* of abstraction. If they had, the training should have facilitated their performance on the far transfer problems as well, but it did not. If the extraction of an abstract schema that biased pattern detection was the main transfer process operating in this scenario, then that schema was encoded at the object level.

This conclusion differs from two alternative claims that have appeared in the literature. On the one hand, our data provide no reason to believe that the knowledge gained during the acquisition phase was limited to surface features, as suggested by Servan-Schreiber and Anderson (1990) and by Perruchet and Gallego (1997) in the context of artificial grammar learning. In the present study, the number of training sequences (twelve) was too small to give the participants an opportunity for statistical learning (e.g., intuitive estimation of bigram frequencies), and the non-overlap between the training and target sequences prevented knowledge of the particular elements in the former from facilitating the processing of the latter. The facilitating effect of our training procedure on the near transfer problems cannot be explained in terms of such surface features.

On the other hand, the formal sequence representations proposed by Greeno and Simon (1974), Restle and Brown (1970) and Simon (1972) hypothesize domain-independent pattern representations that imply transfer across all symbol alphabets (and non-symbolic sequences as well). It is possible that people have the ability to construct pattern representations at that level of abstraction, but our data do not provide any evidence that they spontaneously did so in the present study. In short, the most parsimonious explanation for our data is that the participants engaged in a level of abstraction that is intermediate between the instance level and the level of domain-independent concepts.

Transfer via pattern detection that is biased by a prior schema is quite different from other proposed mechanisms for declarative transfer such as knowledge compilation, constraint violation and analogy. According to the knowledge compilation hypothesis proposed by Anderson (1983) and Neves and Anderson (1981), declarative representations are re-represented in terms of production rules (procedural knowledge). This idea applies naturally to geometry theorems and other declarative knowledge structures with an *if-then* structure, but is not clear what useful procedural knowledge a knowledge compilation mechanism would produce if applied to a simple relational proposition like *33 is two more than 31*. Presumably, the resulting rule would be, approximately, *if you want to get to N_2 and N_2 is two steps further than N_1, and you know how to take two steps, then try to get to N_1*. Such a rule would provide no help in solving sequence extrapolation problems.

According to the constraint violation theory (Ohlsson, 1996), declarative knowledge provides a set of constraints on the correct solution with which the problem solver can identify errors and correct them. This process applies primarily to pattern extrapolation rather than pattern detection. If this process were the main determinant of transfer in our study, then our participants must have generated letters more or less at random, and then thought through whether those letters formed an instance of the abstract pattern. It is logically possible to proceed in this way. However, this generate-and-test approach transforms the extrapolation problem into a search problem with a gigantic search space; with 8 positions and 26 possibilities in each position, the size of the problem space is 26^8, which is an astronomical number. Even if the participants worked very fast, it was not possible to search a space of that size in five minutes via generate and test. It is possible that error correction played an auxiliary role in enabling students to evaluate the letters they generated on the basis of their pattern representation.

A final possibility is that our participants formed structural analogies between the strings encountered during training and the given parts of the extrapolation problems, as structure mapping theories of analogy (Gentner, Holyoak & Kokinov, 2001) would suggest. There are two reasons to doubt that this process played an important role in this case. First, it is not clear how analogy would have helped the participants. They were confronted with a given sequence and asked to extrapolate it. Because the given sequence did in fact share an underlying pattern with the previously encountered training strings, the participants might have mapped the former onto one of the latter (if they were able to recall at least one training string). However, mapping is only the first step in analogical reasoning; the next step is to transfer something from the source to the target. What could the participants have transferred in this case? In order to construct the mapping between the relational structures, the participants would first have had to identify the pattern in the given part of the extrapolation problem. But once the pattern has been identified, the source (the training string) provides no additional information about the given sequence; there is nothing to transfer.

Second, the pattern in the far transfer problems was the same as in the near transfer problems. A structure mapping engine capable of aligning distinct relations of the same order (e. g., *plus 2* and *two steps forward in the alphabet*) should be able to map the structure of the training strings onto the structure of the far transfer problems as easily as onto the near transfer problems, in which case there should have been a facilitating effect for the far transfer problems as well. But we did not observe such an effect. The standard explanation for the difficulty of deep structural analogies – that the right source analogue is not retrieved from memory and hence not available to be mapped -- does not apply to the present case, because the participants were instructed that any pattern in the training strings might be helpful on the problem solving task, a manipulation that is known to strongly affect the probability of source recall (Gick & Holyoak, 1980, 1983).

In short, neither knowledge compilation, constraint violation, nor analogy appear to be a plausible explanation for our results. The simplest and most straightforward explanation is the classical abstraction hypothesis: The participants learned abstract schemas for the relevant patterns during the acquisition phase, and those schemas productively biased their problem perception and attention allocation during the application phase. This is a declarative transfer process.

Were there also procedural transfer processing operating in our study? If we were to explain our results with the identical elements hypothesis (Thorndike & Woodworth, 1901), we would have to claim that the training taught the participants the very same inference rules they needed to extrapolate the given sequence during problem solving. However, the participants did not perform any extrapolation inferences during the acquisition phase; they were not asked to perform any inferences at all. Hence, they had no opportunity to acquire any rules for performing extrapolation inferences, or so it seems.

Procedural transfer might have entered through a back door. Each training string consisted of two complete periods of the underlying pattern, so knowledge of the pattern could be used to improve performance on the recall task. The participants could use what they remembered about one period, plus what they knew about the pattern to generate (rather than recall) parts of the other period, thus turning the recall task into a reasoning task. That is, the participants might have engaged in *reconstructive recall*, and in that way unwittingly practiced making extrapolation inferences. The bimodal distribution of the data for the patterned training groups (see Figure 3, panel a) suggest that the upper 2/3 of the participants approached the task in a different way than the bottom 1/3, so it is possible that the former engaged in reconstructive recall while the latter did not.

There are objections this explanation. First, because the symbol strings used in the extrapolation problems were not the same as those in the training strings, the reconstructive recall rules would have had to be formulated at the object level of abstraction. But there were no incentives for the participants to construct abstract rules during the acquisition phase. Their task in that phase was merely to reproduce each particular sequence. Second, the participants' goal during training was *to commit the symbol strings to memory, to recall the strings accurately* or something similar. During extrapolation their goal presumably was, approximately, *to figure out which letter goes into position N*. In most rule-based systems, rules refer to goals (e.g., Anderson & Lebiere, 1998). The rules acquired during reconstructive recall would not have had the same goals as the extrapolation rules, so they were not identical to the latter and hence could not have transferred in the manner envisioned in the identical rules theory.

These objections can be overcome if we assume that the goals in the rules were represented in a less specific way, e.g., *get the next letter*. This representation applies to both reconstructive recall and to extrapolation. Furthermore, instead of a strict case of identical elements transfer, this might be a case of procedural-to-procedural analogy (Anderson & Thompson, 1989; Carbonell, 1986). Even if the rules that were learned during reconstructive recall were not fully abstract and had a slightly different goal representation, they might nevertheless have served as source analogues or templates for the extrapolation rules. We conclude that an auxiliary role for procedural transfer cannot be ruled out.

In summary, we propose that the two classical hypotheses of abstraction and (almost) identical rules suffice to explain our findings. First, the attempt to commit the training strings to memory led the participants to look for relations within those strings, and to construct a schema for the underlying pattern at the object level of abstraction. The repeated retrieval and application of that schema during training biased perception and attention access in such as way as to enable quick and accurate identification of the pattern in the given strings during problem solving. We suggest that this declarative transfer process was the main determinant of the results we observed.

Second, procedural transfer played an auxiliary role. Some of our participants might have used the relations within the training strings to infer parts of the training strings that they could not recall. Such reconstructive recall served as an opportunity to practice the type of inference that they needed to carry out during extrapolation. The resulting rules might have transferred to the problem solving task because they truly were identical to the needed inference rules, because they were modified by procedural-to-procedural analogy, or because they were encoded at the object level of abstraction even though they did not need to be to function in the recall task.

Implications

The conclusion we reached in the previous section is seemingly at variance with a trend in research on transfer that rejects the traditional concept of transfer as the application of previously learned declarative abstractions to a novel problem or situation (Bransford & Schwartz, 1999; Carraher & Schliemann, 2002; Lave, 1988; Lave & Wenger, 1991). One motivation for this trend is the failure to find transfer in some laboratory paradigms where one would intuitively expect it (e.g., Bassok & Holyoak, 1989; Perfetto, Bransford, & Franks, 1983; Reed, Dempster, & Ettinger, 1985). A second, and more important motivation is the difficulty students appear to have in using abstract declarative knowledge to guide their problem solving in school subjects. For example, conceptual understanding that leads to appropriate problem solving action has long been a highly desired goal of, for example, mathematics instruction but mathematics education researchers have found this program difficult to carry out (e.g., Hiebert, 1986). Even when such understanding is achieved, correct action does not necessarily follow (e.g., Resnick & Omanson, 1987). Finally, the concept of transfer itself (understood as mediated by abstraction) has been criticized as static, ignoring social context, and lacking a motivational component (e.g., Lave, 1988; Carraher & Schliemann, 2002).

The alternative is that what transfers is primarily skills and practices, not theoretical or declarative knowledge. Observations of learning outside deliberately arranged situations like transfer experiments or classrooms support this idea. In many everyday situations, learners acquire a skill or a practice by participating, in a peripheral and preliminary way, in its execution; this is sometimes referred as *apprenticeship learning* (Greeno, 1997a, 1997b; Lave & Wenger, 1991). Many skills and practices acquired in this way are not associated with any abstract theory or set of principles, so declarative transfer plays a minimal role when they are applied in novel contexts. The implication of this perspective is that a training context should be similar to the intended application context. This has indeed long been the case in both athletics (Burton, Brown, & Fischer, 1984), crafts like blacksmithing (Keller & Keller, 1996), the training of naval navigators (Hutchins, 1996), and professional domains from medicine (Patel, Kaufman & Magder, 1996) to pilot training. High-fidelity training environments like flight simulators (Patrick, 1992, chap. 12) can be seen as attempts to ensure the transfer of training via the simplest and most direct transfer process, namely procedural transfer via (almost) identical elements. In general, similarity – indeed, identity – of training and application contexts is a central feature of apprenticeship learning, on-the-job training, and the acquisition of expertise (Ericsson, 1996; Feltovich, Ford & Hoffman, 1997). This is precisely what the principle of transfer-appropriate processing would predict.

However, the prevalence of procedural transfer in everyday contexts does not imply that declarative transfer is non-existent or unimportant. Everyday situations do not always carry their lessons on their sleeves, and we cannot anticipate every future situation or application context. Hence, apprenticeship and prior transfer-appropriate processing cannot always prepare us for future situations. In addition, schooling and academic instruction aim to prepare students for future application contexts which cannot be known beforehand and which necessarily will be different from the classroom or the lecture hall. Several educational practices, such as organizing text book chapters with theory up front and practice problems at the end and lectures followed by independent seat work, implicitly presuppose declarative transfer for their effectiveness. In short, our success at negotiating everyday life would be incomprehensible if we had no ability to transfer declarative knowledge. Consistent with these observations, both the results of prior studies (Cheng, 2002; Carney & Levin, 2000; Rittle-Johnson & Alabali, 1999) and the large transfer effects that we found support to the psychological reality of declarative transfer.

There is no conflict between these perspectives. Declarative and procedural transfer should not be conceived as competing principles of transfer, nor should the various cognitive processes that have been proposed for these two types of transfer be conceived as competing hypotheses. We suggest that a more useful perspective starts with the assumption that the mind possesses a repertoire of potential transfer processes, i.e., processes by which the outcome of prior learning is brought to bear on a future problem or situation. Some (but not necessarily all) of those processes have been identified and described by researchers. Any particular instance of transfer is due to some combination of those processes. The weight of the individual processes in a particular transfer scenario is a function of the specific features of that scenario.

Why did we observe large amounts of declarative transfer in our study? What were the specific features of our scenario that triggered this outcome? First, for declarative transfer to play a role, there must be some declarative knowledge to acquire. In our case, the patterns in the training strings play that role. Second, there must be some opportunity to form an abstract representation of the declarative knowledge. The training phase provided such an opportunity. Third, there must be a clear cognitive function for declarative knowledge in solving the transfer task. In our case, that role is played by the pattern detection component. In contrast, the training phase did not provide a good opportunity to acquire any relevant procedural knowledge (although the participants might possibly have done so anyway, via reconstructive recall). In a scenario with these properties, it is reasonable to expect declarative transfer via abstraction to be the main determinant of the outcome. Other scenarios, with other properties, will no doubt exhibit a different profile or mix of transfer processes. The general implication is that researchers can achieve greater theoretical clarity by paying closer attention to the specific features of transfer scenarios when discussing empirical results.

The key concept in the explanation for our findings is that of an intermediate level of abstraction, i.e., abstraction over objects but not over relations. A simple dichotomy of specific versus abstract leads to the paradox that what was learned during the acquisition phase either was a fully specific representation of the study materials, in which case what was learned could not have transferred even to the near transfer task, or else abstract, in which case what was learned should have transferred to the far transfer task. Neither of these statements is a good description of what we found, nor of what has been found in transfer research generally. It is possible that some contrasting findings in transfer research can be better understood if they are analyzed with respect to which transfer scenarios require abstraction over objects only, and which require abstraction over relations as well.

The notion of an intermediate level of abstraction implies a style of cognition that we might call *bounded flexibility*: Flexible use of knowledge within but not between domains, where a domain is defined in terms of its central relational predicates. This notion seems promising as a tool for understanding not only laboratory results but also the mixture of success and failure in applying what has been learned in one situation to another that characterizes everyday life.

References

Adams, L. T., Kasserman, J. E., Yearwood, A. A., Perfetto, G. A., Bransford, J. D., & Franks, J. J. (1988). Memory access: The effects of fact-oriented versus problem-oriented acquisition. *Memory & Cognition, 16*, 167-175.

Altmann, G. T. M., Dienes, A., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 899-912.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Greeno, J. G., Kline, P. J., & Neves, D. M. (1981). Acquisition of problem-solving skill. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 191-230). Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

Anderson, J. R., & Thompson, R. (1989). Use of analogy in a production-system architecture. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 267-297). Cambridge, England: Cambridge University Press.

Bassok, M. (1990). Transfer of domain-specific problem-solving procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 3*, 522-533.

Bassok, M., & Holyoak, K. J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 153-166.

Beck, I. L., & McKeown, M. G. (1983). Learning words well – A program to enhance vocabulary and comprehension. *Reading Teacher, 36*, 622-625.

Berry, D., (Ed.), (1997). *How implicit is implicit learning?* Oxford, UK: Oxford University Press.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education, 24*, 61-100.

Bransford, J. D., Sherwood, R., Vye, N., & Reiser, J. (1986). Teaching thinking and problem solving. *American Psychologist, 4*, 1078-1089.

Brown, A. L., & Kane, M. J. (1988). Preschool children can learn to transfer: Learning to learn and learning from example. *Cognitive Psychology, 20*, 493-523.

Burton, R., Brown, J. S., & Fischer, G. (1984). Skiing as a model of instruction. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 139-150). Cambridge, MA: Harvard University Press.

Carbonell, J. G. (1986). Derivational analogy: A theory of reconstructive problem solving and expertise acquisition. In R. S. Michalski, J. G. Carbonell & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (Vol. 3, pp. 371-392). Los Altos, CA: Kaufmann.

Carney, R. N., & Levin, J. R. (2000). Mnemonic instruction, with a focus on transfer. *Journal of Educational Psychology, 92*, 783-790.

Carraher, D., & Schliemann, A. D. (2002). The transfer dilemma. *The Journal of the Learning Sciences, 11*, 1-24.

Catrambone, R. (2002). The effects of surface and structural feature matches on the access of story analogs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 318-334.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1147-1156.

Chen, Z. (2002). Analogical problem solving: A hierarchical analysis of procedural similarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 81-98.

Chen, Z., & Daehler, M. W. (1989). Positive and negative transfer in analogical problem solving by six-year-old children. *Cognitive Development, 4*, 327-344.

Cheng, P.C. H. (2002). Electrifying diagrams for learning: Principles for complex representational systems. *Cognitive Science, 26*, 685-736.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Chi, M. T. H., Slotta, J. D., & deLeeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction, 4*, 27-43.

Clement, J. (1982). Algebra word problem solutions: Thought processes underlying a common misconception. *Journal for Research in Mathematics Education, 13*, 16-30.

Cohen, J. (1988). *Statistical power analysis of the behavioral sciences.* (2nd Ed.) New York: Academic Press.

Cormier, S. M., & Hagman, J. D., (Eds.), (1987). *Transfer of learning: Contemporary research and applications.* New York: Academic Press.

Cognition and Technology Group at Vanderbilt (1994). From visual word problems to learning communities: Changing conceptions of cognitive research. In K. McGilly (Ed.). *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 157-200). Cambridge, MA: MIT Press.

Detterman, D. K., & Sternberg, R. J., (Eds.) (1993). *Transfer on trial: intelligence, cognition, and instruction.* Norwood, NJ: Ablex.

Ericsson, K. A., (Ed.), (1996). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports and games.* Mahwah, NJ: Erlbaum.

Feltovich, P., Ford, K., & Hoffman, R., (Eds.), (1997). *Expertise in context: Human and machine.* Menlo Park, CA: AAAI Press.

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*, 141-205

Gagne, R. M. (1962). The acquisition of knowledge. *Psychological Review, 69*, 355-365.

Gagne, R. M. (1965). *The conditions of learning.* New York: Holt, Rinehart, & Winston.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teaming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 99-129). Hillsdale, NJ: Erlbaum.

Gentner, D., Holyoak, K. J., & Kokinov, B. N., (Eds.), (2001). *The analogical mind.* Cambridge, MA: MIT Press.

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.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277-300.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Gick, M. L., & Holyoak, K. J. (1987). The cognitive basis of knowledge transfer. In S. M. Cormier & J. D. Hagman (Eds.), *Transfer of learning: Contemporary research and applications* (pp. 9-46). New York: Academic Press.

Goldstone, R. L., & Sakamoto, Y. (in press). The transfer of abstract principles governing adaptive systems. *Cognitive Psychology.*

Goodman, N. (1972). *Problems and projects.* New York, NY: Bobbs-Merill.

Grenno, J. G., (1997a). Response: On claims that answer the wrong questions. *Educational Researcher, 26*, 5-17.

Grenno, J. G. (1997b). Theories and practices in thinking and learning to think. *American Journal of Education, 106*, 85-126.

Greeno, J. G., & Simon, H. A. (1974). Processes for sequence production. *Psychological Review, 81*, 187-198.

Hiebert, J., (Ed.), (1986). *Conceptual and procedural knowledge: The case of mathematics.* Hillsdale, NJ: Erlbaum.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition, 15*, 332-340.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295-355.

Holyoak, K. J., Novick, L. R., & Melz, E. R. (1994). Component processes in analogical transfer: Mapping, pattern completion, and adaptation. In K. J. Holyoak & J. A. Barden (Eds.), *Advances in connectionist and neural computation theory volume 2: Analogical connections* (pp.113-180). Norwood, NJ: Ablex Publishing Company.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427-466.

Hutchins, E. (1996). Learning to navigate. In S. Chaiklin & J. Lave (Eds.), *Understanding practice* (pp. 35-63). Cambridge, UK: Cambridge University Press.

Judd, C. H. (1908). The relation of special training to general intelligence. *Educational Review, 36*, 28-42.

Karmiloff-Smith, A. (1992). *Beyond modularity.* Cambridge, MA: MIT Press.

Katona, G. (1940). *Organizing and memorizing.* New York: Columbia University Press.

Keane, M. T. (1987). On retrieving analogues when solving problems. *Quarterly Journal of Experimental Psychology, 39A*, 29-41.

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

Keller, C.M., & Keller, J. D. (1996). *Cognition and tool use: The blacksmith at work.* Cambridge, UK: Cambridge University Press.

Klahr, D. & Carver, S. M. (1988). Cognitive objectives in a LOGO debugging curriculum: Instruction, learning, and transfer. *Cognitive Psychology, 20*, 362-404.

Klahr, D., & Wallace, J. G. (1970). An information processing analysis of some Piagetian experimental tasks. *Cognitive Psychology, 1*, 350-387.

Kieras, D. E., & Bovair, S. (1986). The acquisition of procedures from text: A production system analysis of transfer of training. *Journal of Memory and Language, 25*, 507-524.

Kotovsky, K., & Simon, H. (1973). Empirical tests of a theory of human acquisition of concepts for sequential patterns. *Cognitive Psychology, 4*, 399-424.

Lave, J. (1988). *Cognition in practice.* Cambridge, MA: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge, MA: Cambridge University Press.

Lee, A. Y., & Pennington, N. (1993.) Learning computer programming: A route to general reasoning skills? In C. R. Cook, J. C. Scholtz & J. C. Spoher (Eds.), *Empirical studies of programmers: Fifth workshop* (pp. 113-136). Norwood, NJ: Ablex.

Lockhart, R. S., Lamon, M., & Gick, M. L. (1988). Conceptual transfer in simple insight problems. *Memory & Cognition, 16*, 36-44.

Manza, L., & Reber, A. S. (1997) Representing artificial grammars: Transfer across stimulus forms and modalities. In D. C. Berry (Ed.), *How implicit is implicit learning?* (pp. 73-106). Oxford, UK: Oxford University Press.

Marshall, S. P. (1995). *Schemas in problem solving*. Cambridge, UK: Cambridge University Press.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254-278.

Mitrovic, A., & Ohlsson, S. (1999). Evaluation of a constraint-based tutor for a database language. *International Journal of Artificial Intelligence and Education, 10*, 238-256.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519-533.

National Research Council (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Neves, D. M., & Anderson, J. R. (1981). Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 57-84). Hillsdale, NJ: Erlbaum.

Novick, L. (1988). Analogical transfer, problem similarity and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 510-520.

Ohlsson, S. (1993). Abstract schemas. *Educational Psychologist, 28*, 51-66.

Ohlsson, S. (1996). Learning from performance errors. *Psychological Review, 103*, 241-262.

Ohlsson, S., Ernst, A., & Rees, E. (1992) The cognitive complexity of doing and learning arithmetic. *Journal for Research in Mathematics Education, 23*, 441-467.

Ohlsson, S., & Lehtinen, E. (1997). The role of abstraction in the assembly of complex ideas. *International Journal of Educational Research, 27*, 37-48.

Ohlsson, S., & Rees, E. (1991) The function of conceptual understanding in the learning of arithmetic procedures. *Cognition & Instruction, 8*, 103-179.

Ohlsson, S., & Regan, S. (2001). A function for abstract ideas in conceptual learning and discovery. *Cognitive Science Quarterly, 1*, 243-277.

Patel, V., Kaufman, D., & Magder, S. (1996). The acquisition of medical expertise in complex dynamic environments. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 127-165). Mahwah, NJ: Erlbaum.

Patrick, J. (1992). *Training: Research and practice*. London, UK: Academic Press.

Perfetto, G. A., Bransford, J. D., & Franks, J. J. (1983). Constraints on access in a problem solving context. *Memory & Cognition, 11*, 24-31.

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher, 17*, 16-25.

Perruchet, P. & Gallego, J. (1997) A subjective unit formation account of implicit learning. In D. C. Berry, (Ed.), *How implicit is implicit learning?* (pp. 124-161). Oxford, UK: Oxford University Press.

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development, 6*, 449-468.

Postman, L. (1964). Short-term memory and incidental learning. In A. W. Melton, (Ed.), *Categories of Human Learning* (pp. 145-201). New York, NY: Academic Press Inc.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 317-327.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219-235.

Reber, A. S. (1993) *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York, NY: Oxford University Press.

Reed, S. K. (1987). A structure mapping model for word problems. *Journal of experimental Psychology: Learning, Memory, & Cognition, 13*, 127-139.

Reed, S. K. (1993). A schema-based theory of transfer. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 39-67). Norwood, New Jersey: Ablex Publishing Company.

Reed, S. K., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 106-125.

Resnick, L. B., & Omanson, S. F. (1987). Learning to understand arithmetic. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 3, pp. 41-95). Hillsdale, NJ: Erlbaum.

Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review, 77*, 481-495.

Restle, F., & Brown, E. (1970). Organization of serial pattern learning. In G. Bower, (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 4, pp. 249-331). New York: Academic Press.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39*, 475-543.

Rittle-Johnson, B., & Alabali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*, 175-189.

Ross, B. H., & Kilbane, M. C. (1997). Effects of principle explanation and superficial similarity on analogical mapping in problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*, 427-440.

Royer, J. M. (1979). Theories of transfer of learning. *Educational Psychologist, 14*, 53-69.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. Bobrow & A. Collins (Eds.), *Representation and understanding* (pp. 211-236). New York: Academic Press.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomena. *Educational Psychologist, 24*, 113-142.

Salvucci, D. D., & Anderson, J. R. (2001). Integrating analogical mapping and general problem solving: the path mapping theory. *Cognitive Science, 25*, 67-110.

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 592-608.

Schwartz, D. L., Lin, X. D., Brophy, S., & Bransford, J. D., (1999). Toward the development of flexibly adaptive instructional design. In C. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm for instructional theory* (pp. 183-213). Mahwah, NJ: Erlbaum.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 501-518.

Schank, R. (1986). *Explanation patterns*. Hillsdale, NJ: Erlbaum.

Shoenfeld, A. H. (1985). *Mathematical Problem Solving*. Orlando, FL: Academic Press.

Simon, H. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review, 79*, 369-382.
Simon, H. A., & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review, 70*, 534-546.
Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill.* Cambridge, MA: Harvard University Press.
Stadler, M., & Frensch, P, (Eds.), (1998). *Handbook of implicit learning.* Thousand Oaks, CA: SAGE.
Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8*, 247-261.
Thorndyke, P. W. (1984). Applications of schema theory in cognitive research. In J. R. Anderson & S. M. Kosslyn (Eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower* (pp. 167-192). New York, NY: Freeman & Company.
Thurstone, L., & Thurstone, T. (1941). *Factorial studies in intelligence.* Chicago: University of Chicago Press.
Wattenmaker, W. D. (1999). The influence of prior knowledge in intentional versus incidental concept learning. *Memory & Cognition, 27*, 685-698.
Wertheimer, M. (1945). *Productive Thinking.* New York: Harper & Row.

Appendix

Twelve Patterned Training Strings for Problem 1

Example	String					
1	13 15	35 16	14 37	16 18	35 19	17 37
2	21 23	43 24	22 45	24 26	43 27	25 45
3	59 61	81 62	60 83	62 64	81 65	63 83
4	47 49	69 50	48 71	50 52	69 53	51 71
5	30 32	52 33	31 54	33 35	52 36	34 54
6	42 44	64 45	43 66	45 47	64 48	46 66
7	63 65	85 66	64 87	66 68	85 69	67 87
8	17 19	39 20	18 41	20 22	39 23	21 41
9	38 40	60 41	39 62	41 43	60 44	42 62
10	52 54	74 55	53 76	55 57	74 58	56 76
11	71 73	93 74	72 95	74 76	93 77	75 95
12	44 46	66 47	45 68	47 49	66 50	48 68

Twelve Patterned Training Strings for Problem 2

Example	String					
1	26 27	25 28	24 36	27 28	26 29	25 37
2	13 14	12 15	11 23	14 15	13 16	12 24
3	40 41	39 42	38 50	41 42	40 43	39 51
4	51 52	50 53	49 61	52 53	51 54	50 62
5	83 84	82 85	81 93	84 85	83 86	82 94
6	77 78	76 79	75 87	78 79	77 80	76 88
7	17 18	16 19	15 27	18 19	17 20	16 28
8	56 57	55 58	54 66	57 58	56 59	55 67
9	29 30	28 31	27 39	30 31	29 32	28 40
10	61 62	60 63	59 71	62 63	61 64	60 72
11	45 46	44 47	43 55	46 47	45 48	44 56
12	85 86	84 87	83 95	86 87	85 88	84 96

Rendering Information Processing Models of Cognition and Affect Computationally Explicit: Distributed Executive Control and the Deployment of Attention[1]

Philip J. Barnard
MRC Cognition and Brain Sciences Unit

Howard Bowman[2]
University of Kent at Canterbury

In this paper we illustrate the potential of process algebra to implement modular mental architectures of wide scope in which control is distributed rather than centralised. Drawing on the Interacting Cognitive Subsystems (ICS) mental architecture, we present an implemented model of the attentional blink effect. The model relies on process exchanges between propositional meaning and a more abstract, implicational level of meaning, at which affect is represented and experienced. We also discuss how the proposed mechanism of buffer movement can, in the context of the ICS architecture, be extended to account for effects of emotional stimuli and brain damage.

Keywords: Distributed systems, process algebra, executive control, attentional blink, emotion, interacting cognitive subsystems

Introduction

Production system architectures and connectionism have dominated research on computational realisation of theories of cognition. However, in spite of great technical progress in sub-symbolic, symbolic and hybrid methods, much influential theory still lacks computational realisation and interactions between processes, subsystems or mental modules are still often

[1] Acknowledgments: We would like to thank the reviewers of this paper who made valuable suggestions, which have greatly improved the presentation of our ideas.

[2] Address for correspondence: Computing Laboratory, University of Kent at Canterbury, Canterbury, Kent, UK, CT1 2DE (Email: H.Bowman@ukc.ac.uk, Telephone: +44 1227 823815, Fax: +44 1227 762811).