Psychological Review
2002, Vol. 109, No. 2, 401–448

0033-295X/02/$5.00 DOI: 10.1037//0033-295X.109.2.401

Theoretical and Computational Analysis of Skill Learning, Repetition Priming, and Procedural Memory

Prahlad Gupta
University of Iowa

Neal J. Cohen
University of Illinois at Urbana-Champaign

This article analyzes the relationship between skill learning and repetition priming, 2 implicit memory phenomena. A number of reports have suggested that skill learning and repetition priming can be dissociated from each other and are therefore based on different mechanisms. The authors present a theoretical analysis showing that previous results cannot be regarded as evidence of a processing dissociation between skill learning and repetition priming. The authors also present a single-mechanism computational model that simulates a specific experimental task and exhibits both skill learning and repetition priming, as well as a number of apparent dissociations between these measures. These theoretical and computational analyses provide complementary evidence that skill learning and repetition priming are aspects of a single underlying mechanism that has the characteristics of procedural memory.

One of the most significant developments in the study of human memory over the last two decades has been the discovery of a dissociation between two different kinds of memory systems. An early indication of this dissociation came from studies of amnesia in patients with excision or lesions of the hippocampus (Scoville & Milner, 1957). These patients were dramatically impaired in their ability to recollect new events and experiences. Cohen and Squire (1980) introduced the term *declarative memory* to refer to this kind of memory, and it is now well accepted that such memory relies on the hippocampus and related medial temporal lobe structures, and that it is profoundly impaired in patients with amnesia (see Cohen & Eichenbaum, 1993, for review). Impairments in this memory system can be revealed by direct tests of memory, which require explicit retrieval of the contents of specific experiences; a typical example would be memory for arbitrary pairings of words. Indication of a second memory system came in the form of a striking pattern of preservation of certain abilities in patients with amnesia (Cohen & Squire, 1980). The forms of memory and learning that are spared by hippocampal damage include the acquisition of skills that are acquired gradually over several sessions of practice; and facilitation (priming) in processing of a stimulus, following prior exposure to that stimulus. Patients with amnesia exhibit normal patterns of skill acquisition and priming in such tasks, provided there is no requirement for direct recollection of previous exposure or practice (for review, see Schacter, Chiu, & Ochsner, 1993). This second kind of learning and memory has been termed *implicit memory* (Schacter, 1987), *nondeclarative memory* (Squire, 1992), or *procedural memory* (Cohen & Eichenbaum, 1993), and it is examined by indirect tests, which do not require explicit recollection of previous experiences.

Early findings from patients with amnesia were followed by the discovery that dissociations between declarative and procedural memory could be obtained in normal populations (e.g., Graf & Schacter, 1985). These results generated much interest in and a large body of research on “implicit” and “explicit” memory or (to use the terms we prefer) on declarative and procedural memory (for review, see Roediger & McDermott, 1993). Much of the interest in procedural (implicit) memory arises from the fact that a great deal of everyday human learning appears to have just the character of procedural memory: It occurs gradually, as a result of practice over many exposures; and the results or contents of such memory, learning, or knowledge are typically unavailable to introspection or recollection. For example, learning to ride a bicycle appears to have these characteristics, as does acquisition of the phonology or syntax of a first language. Thus the study of procedural memory makes contact with rich traditions of psychological inquiry, such as the investigation of skill acquisition, and investigation of language learning. Further specification of the nature and mechanisms of procedural memory therefore holds the promise of providing fresh insight into fundamental and pervasive human learning processes.

The aim of this article is to make a systematic examination of the relationship between skill learning and repetition priming, with a view to obtaining fresh insight into the nature of procedural memory. Cohen (1984; Cohen & Eichenbaum, 1993) put forth the

Prahlad Gupta, Department of Psychology, University of Iowa; Neal J. Cohen, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign.

A brief discussion of one of the simulations reported here was presented at the 28th Carnegie Mellon Symposium on Cognition, Pittsburgh, Pennsylvania, 1997.

We acknowledge our indebtedness to the work of Russ Poldrack and Scott Selco (Poldrack, 1995; Poldrack et al., 1999), which provided the paradigm and the motivation for us to explore the issues discussed here. We thank Gary Dell, for invaluable input of many kinds, and also Kay Bock, Barbara Church, Cindy Fisher, Susan Garnsey, Gregory Murphy, Courtney Nash, Brian Ross, and Scott Selco for helpful discussion of some of the ideas presented here. We also thank Gordon Logan for comments he provided as a colleague at the University of Illinois.

Correspondence concerning this article should be addressed to Prahlad Gupta, Department of Psychology, University of Iowa, Iowa City, Iowa 52242. E-mail: prahlad-gupta@uiowa.edu

hypothesis that a number of apparently heterogeneous implicit memory phenomena, such as skill learning, repetition priming, and habit formation, are subserved by a single kind of learning mechanism termed *procedural memory*, which is based on the continual, experience-driven tuning of processing elements. This hypothesis does not claim that there is a single or specific set of processing elements that constitute a procedural memory "system." Rather, it claims that there is a single kind of procedural learning mechanism namely, incremental tuning of the processing elements that underlie a particular task, whatever or wherever those processing elements may be. This tuning can be thought of as the kind of adjustment of connection weights that occurs in parallel distributed processing (PDP) networks (McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986); a similar view of procedural memory is inherent in the work of some of the architects of the PDP framework (McClelland, McNaughton, & O'Reilly, 1995; McClelland & Rumelhart, 1985). According to this hypothesis, then, skill learning and repetition priming are both manifestations of procedural memory. However, there have also been a number of reports that suggest that skill learning and repetition priming can be dissociated from each other and, consequently, that there are different mechanisms that underlie these two implicit memory phenomena (e.g., Heindel, Salmon, Shults, Wallicke, & Butters, 1989; Kirsner & Speelman, 1996; Schwartz & Hashtroudi, 1991). The status of this debate is currently unresolved and is a matter of some contention. If these phenomena are indeed based on different underlying mechanisms, then they cannot both be regarded as aspects of a single procedural memory mechanism. On the other hand, if skill learning and repetition priming can be shown to arise from a single underlying mechanism, then the hypothesis that they are both forms of procedural memory is strengthened. The question of whether skill learning and repetition priming arise from the same or from different mechanisms thus has important consequences for thinking about the nature of procedural memory as a whole.

Repetition priming refers to facilitation (as seen in greater accuracy or faster performance) in the processing of specific items (stimuli) in a task as a result of previous exposure to those items. In the implicit memory literature, *skill learning* refers to task performance improvement that is not restricted to specific items to which there has been previous exposure, but that extends to new items; it is the development of generalized task ability as a result of practice in the task. Although there is evidence from both neuropsychologically impaired and nonimpaired populations to suggest that skill learning and repetition priming have various features in common (e.g., Cohen & Squire, 1980; Logan, 1990), dissociations between these phenomena have also been observed, and these dissociations have been interpreted as indicating that the phenomena of skill learning and repetition priming may arise from different underlying mechanisms. For example, Logan (1990) reported that skill learning and repetition priming have three characteristics in common. First, they both appear to increase as a power function of the number of exposures. Second, they both appear to be stimulus specific. Third, they both appear to depend on associations between stimulus and responses. These results suggested that skill learning and repetition priming might arise from the same underlying mechanism. A contrary suggestion was made by Schwartz and Hashtroudi (1991), who examined repetition priming and skill learning in each of three tasks: partial-word identification, inverted reading, and word-fragment completion (Schwartz & Hashtroudi, 1991, Experiment 1). They found that, whereas the magnitude of repetition priming was similar across these tasks, the pattern of skill learning varied, with skill learning occurring in only two of the tasks. They suggested that these divergent patterns indicate that skill learning and repetition priming may arise from different underlying mechanisms. They also found that priming effects were not correlated with improvements in skill across trials, in either the partial-word identification or inverted reading tasks, and suggested that these results provide further indication of a separation of the mechanisms underlying skill learning and repetition priming. Further challenges to the view that skill learning and repetition priming arise from a single mechanism have come from a recent report (Kirsner & Speelman, 1996), which found that, unlike skill learning, repetition priming did not follow the power law of practice in a lexical decision task. Kirsner and Speelman also noted that these results appear to directly contradict Logan, who found that repetition priming does follow the power law of practice in a lexical decision task.

In this article, we present a new approach to thinking about skill learning and repetition priming. We begin by making a functional analysis of the phenomena of skill learning and repetition priming, asking what kinds of relationships might be observed between them and what these various relationships might imply about the underlying processing. In light of this theoretical framework, we reanalyze data that have been taken as evidence for separate mechanisms and argue that none of these data are valid as evidence of a processing dissociation. The second part of the article begins by outlining our view of the mechanism that underlies procedural memory and learning. We discuss how, according to this framework, skill learning and repetition priming are manifestations of the operation of a single underlying learning mechanism. We then present a computational model and simulations of a digit-entering task in which both skill learning and repetition priming have been studied. To our knowledge, this work is the first computational investigation of skill learning and repetition priming in a specific empirical domain. We show that the model provides an accurate account of behavioral data from the digit-entering task. In particular, the model exhibits both skill learning and repetition priming and also exhibits a number of the dissociations between these measures that have been taken as evidence for dual mechanisms. All of these effects are exhibited even though the model consists of only a single learning mechanism.

We view these two parts of the article as providing complementary evidence that skill learning and repetition priming are aspects of a single underlying mechanism. The first part suggests that the interpretation of existing data as evidence of separate mechanisms is unwarranted. The second part of the article articulates our theory of procedural memory and describes a computational model that incorporates this theory. The model provides a computational demonstration of how skill learning, repetition priming, and patterns of apparent dissociation between them can all arise from a single mechanism. We believe that these demonstrations provide strong evidence that skill learning and repetition priming reflect a single underlying learning mechanism and thus support the hypothesis that they are both forms of procedural memory. Additionally, we believe that the arguments presented here help to clarify the terms of the current debate about skill learning and repetition priming and thus constitute one step toward its resolution.

In the concluding section of the article we use our theoretical framework to analyze the conditions under which and the levels of processing at which skill learning and repetition priming can be expected to arise. We also discuss the broader implications of our processing framework, relating it to debates over the nature of memory and to the domains of skill acquisition, automaticity, and language processing and learning. In the remainder of this introductory section, we describe an example of the kind of task in which skill learning and repetition priming can be studied.

A Hypothetical Task

To provide the reader with a feel for the kinds of experimental setting in which both skill learning and repetition priming can be studied, let us consider a hypothetical task. In this hypothetical task, drawings of physically impossible objects are presented to participants on a computer visual display. (The objects are impossible because they contain mutually inconsistent surfaces of the kind featured, for instance, in the art of M. C. Escher). Each of the pictured objects is impossible in one, two, three, or four ways. At each stimulus presentation, the participant's task is to press one of four response keys (representing the four levels of impossibility) to indicate in how many ways the drawing is impossible. Participants' response times are recorded.

Stimulus presentation is divided into blocks of 20 stimuli, although there is no pause between blocks. There are two kinds of stimuli in each block, which we will refer to as *unique stimuli* and *repeating stimuli*. A unique stimulus is presented once in a particular block and does not appear at any other point in the experiment. A repeating stimulus appears once in every block and thus repeats from block to block throughout the experiment. Each block comprises presentation of 10 unique stimuli and 10 repeating stimuli (the latter remaining constant throughout the experiment). The experiment consists of 12 blocks overall. It thus has the structure of a multiple-repetition paradigm of the kind commonly used in studies of skill learning and repetition priming (e.g., Cohen & Squire, 1980; Kirsner & Speelman, 1996; Schwartz & Hashtroudi, 1991).

Figure 1 depicts hypothetical results from this hypothetical task, measured by (hypothetical) mean reaction times. The upper curve in the figure plots performance on unique stimuli across blocks, and the lower curve plots performance on repeating stimuli across blocks. The upper curve depicts substantial improvement in performance on unique stimuli across blocks of practice. Because the unique stimuli are drawings that participants have never previously viewed, the improvement in performance on these stimuli reflects a generalized facilitation in performance of the task, a facilitation that is not restricted to specific stimuli, and extends to novel stimuli. This generalized facilitation is what is termed skill learning. The improvement in performance on unique stimuli is thus a manifestation of the existence of skill learning in this hypothetical task.

The lower curve in the figure shows performance on repeating stimuli improving even more substantially across blocks than performance on unique stimuli. This greater improvement in performance on repeating stimuli indicates facilitation that is due to the repetition of those specific stimuli; this is what is termed repetition priming. Thus, the greater improvement in performance

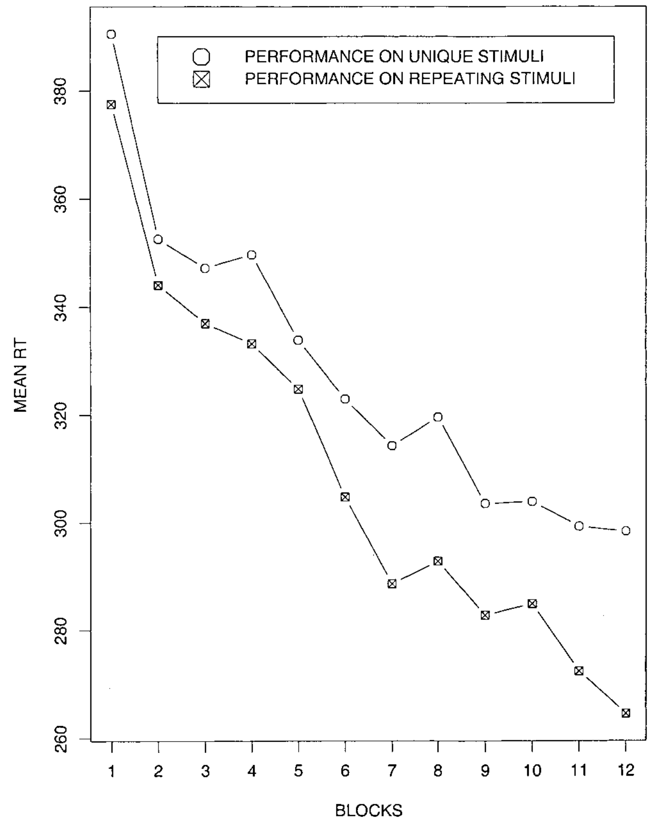


Figure 1. Hypothetical results from a hypothetical object-impossibility judgment task. RT = reaction time.

on repeating stimuli than on unique stimuli is a manifestation of the existence of repetition priming in this hypothetical task.[1]

Theoretical Analysis

In developing our theoretical analysis, we will focus on three issues with respect to which the relationship between skill learning and repetition priming has been studied (e.g., Logan, 1990; Kirsner & Speelman, 1996; Schwartz & Hashtroudi, 1991). The first of these issues relates to the form of functions that may be expected for skill learning and repetition priming. The underlying intuition here has been that, if skill learning and repetition priming are related, then the two functions should exhibit the same shape. This idea has been most explicit in the work of Logan (1990) and Kirsner and Speelman (1993, 1996). These investigators reasoned that repetition priming should exhibit a power-function speedup with practice of the kind that is well established for skill learning (Anderson, 1982; Newell & Rosenbloom, 1981). Logan reported the existence of a power function for repetition priming in a lexical

[1] Note that our hypothetical task is intended merely to illustrate the nature of the skill learning and repetition priming phenomena under consideration, the nature of relationships between them, and the structure of the kinds of paradigms used in many studies. Our hypothetical results are not intended as precise predictions of what might actually obtain if this impossibility-judgment experiment were conducted.

decision task and viewed this as evidence that skill learning and repetition priming are related. However, Kirsner and Speelman reported that repetition priming in a lexical decision task did not follow a power function and in fact did not even increase with practice; they took this to indicate that skill learning and repetition priming may arise from different mechanisms. Thus, each of these investigators drew conclusions about the relationship between skill learning and repetition priming that were based on whether they obtained the predicted power function speedups. As we shall see, some of the apparent conflict in these results turns out to be no more than an artifact of differences in terminology.

A second issue relates to correlations between skill learning and repetition priming. The underlying intuition here is that, if skill learning and repetition priming are related, then they should also be correlated. The correlations here refer to correlations across participants at a particular point in the experiment. This idea is most apparent in the work of Schwartz and Hashtroudi (1991). These authors drew conclusions about the relationship between skill learning and repetition priming that were based on their pattern of correlation across participants at a particular trial in three different experimental tasks. We will show that a lack of correlation between the standard measures of skill learning and repetition priming is simply an artifact of the nature of the measures and does not warrant inferences about the nature of the underlying mechanisms.

A third issue concerns patterns of increase and decrease in skill learning and repetition priming: whether the measures exhibit the same pattern of increase or decrease over the course of experimental practice. Note that this question amounts to asking whether skill learning and repetition priming are correlated over the course of an entire experiment (rather than across participants at a particular point in the experiment). The underlying idea is that, if skill learning and repetition priming are related, then they should both increase or decrease together over the course of experimental trials. This intuition is also exemplified by the work of Schwartz and Hashtroudi (1991), who found that repetition priming increased with practice in some experimental tasks in which there was no evidence for skill learning. The authors interpreted this finding as suggesting that there may be different mechanisms that underlie skill learning and repetition priming. We will show that such patterns of dissociation are consistent with the existence of a single underlying mechanism.

The theoretical framework we develop in the next few sections will focus primarily on these three issues. We will then use our theoretical framework to reanalyze several apparently contradictory empirical results.

Framework and Terminology

We begin development of our analysis by presenting an organizing framework and some terminology. This framework formalizes and extends discussion of the measures that we have already introduced in the context of our hypothetical object-impossibility judgment task. We will continue to use those hypothetical results as a basis for thinking about the relationship between skill learning, repetition priming, and other measures to emphasize that the relationships illustrated are general and apply to any task in which both skill learning and repetition priming can be observed. Figure 2a is a schematic of results from that hypothetical task and illustrates several related measures.

Performance on unique stimuli is a measure of participants' performance in response to presentation of unique stimuli, that is, the stimuli that vary from block to block. Performance is measured in terms of some index of reaction time. For example, in a lexical decision task (e.g., Kirsner & Speelman, 1996), this might be a measure of lexical decision latency for those stimuli that were unique to each block.

Performance on repeating stimuli is a measure of participants' performance in response to presentation of repeating stimuli, that is, those stimuli that appear on every block. This is also measured in terms of reaction time.

Skill learning refers to the development of a generalized ability (i.e., skill), which is not specific to particular stimuli but extends to new items (as long as these items are from the same sampling population). It is measured by improvement in participants' performance on unique stimuli. Thus, skill learning at any given block *n* is measured as Block 1 performance on unique stimuli less Block *n* performance on unique stimuli. The skill learning measure is depicted in two ways in Figure 2a: first, as the divergence of Block *n* performance on unique stimuli from Block 1 performance on unique stimuli; and second, with the same divergence information graphed as a function in the lower part of the figure.

It is important to clearly distinguish between performance on unique stimuli and skill learning. Performance on unique stimuli is a direct measure of how participants perform on novel items (drawn from the sampling population). Skill learning measures the improvement in performance on novel ("unique") stimuli; it is derived from performance on unique stimuli.

Benefit of repetition is a measure of improvement in participants' performance on repeating stimuli. It is depicted in two ways in Figure 2a: first, as the divergence of Block *n* performance on repeating stimuli from Block 1 performance on repeating stimuli; and second, with the same divergence information graphed as a function in the lower part of the figure.

We define the measure of *repetition priming* to be the difference between performance on unique stimuli and performance on repeating stimuli. Figure 2a depicts the repetition priming measure in two ways: first, as the separation between performance on unique stimuli and performance on repeating stimuli, and second, with this same information graphed in the lower part of the figure.

We have already emphasized the distinction between performance on unique stimuli and skill learning. In additional, the measure that we term repetition priming needs to be carefully distinguished from performance on repeating stimuli; and from what we term benefit of repetition, which, as described above, measures the improvement in performance on repeating stimuli.

These terminological points are not trivial, because repetition priming has been used differently by different researchers and in different senses. Schwartz and Hashtroudi (1991) define repetition priming as "facilitation in the processing of an item as a result of previous exposure to the same item" (p. 1177). The measure of priming that Schwartz and Hashtroudi used was "obtained by subtracting the proportion of non-repeated (new) items identified from the proportion of repeated (old) items identified on each trial" (p. 1180). This corresponds to the measure of repetition priming we defined above. Similarly, Logan (1990) defines repetition priming as follows: "Responses are usually faster on the second

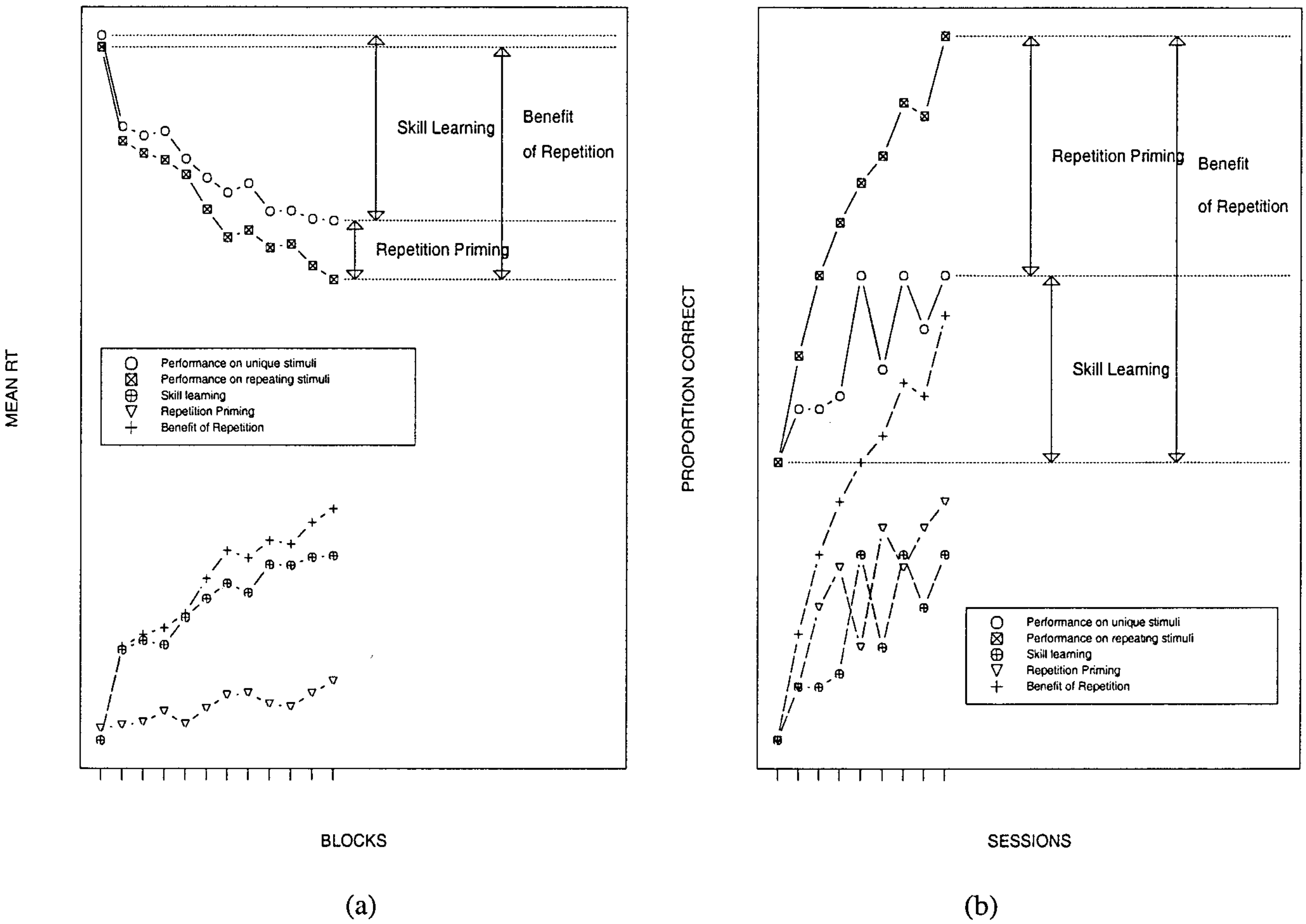


Figure 2. Schematic of relevant measures. (a) Measured by reaction time, in a hypothetical task. (b) Measured by accuracy. Schematic of data from Schwartz and Hashtroudi (1991, Experiment 1, partial-word identification). RT = reaction time.

presentation [of an item] than on the first, and this difference is called *repetition priming*" (p. 2). To measure repetition priming, Logan uses "the reaction time data . . . presented as *benefit scores* . . ." (p. 10), and notes that "benefit . . . was calculated by subtracting the mean reaction time for each number of presentations from the mean reaction time for the first presentation" (p. 10). Note that this measure is equivalent to what we have termed benefit of repetition. Kirsner and Speelman (1996, p. 570) measure repetition priming by the difference between reaction time for new items and reaction time for old items, in each block. This corresponds to the measure of repetition priming that we defined above. Although these comparisons might appear digressive, the importance of terminology will become clear very shortly.

Before moving on to our analysis of the relationship between the various measures identified above, we need to discuss Figure 2b. This figure is a schematic derived from the data of Schwartz and Hashtroudi (1991, Experiment 1, partial-word identification). It illustrates the case in which performance is measured by some criterion of accuracy rather than by reaction time. In consequence, improvement in performance is reflected in performance on unique stimuli and performance on repeating stimuli functions that increase with practice. This can be compared with Figure 2a, where improvement in performance is reflected by decreasing performance on unique stimuli and performance on repeating stimuli reaction time functions. Note, however, that improvement in performance for the other measures (skill learning, repetition priming, and benefit of repetition) is reflected in increasing functions in both figures, that is, irrespective of whether the performance on unique stimuli and performance on repeating stimuli functions denote reaction time (and hence are decreasing functions) or accuracy (and hence are increasing functions). This follows from the fact that skill learning, repetition priming, and benefit of repetition are difference measures. Finally, it should be clear that our observations regarding distinctions between the five measures apply equally, whether performance is measured by reaction time or by accuracy.

The Form of Functions

In this section, we examine the first of the issues we highlighted: What form should the functions for skill learning and repetition priming be expected to take? We will address this question by analyzing the form of functions for all five of the measures we described in the preceding discussion: performance on unique stimuli, performance on repeating stimuli, benefit of repetition, skill learning, and repetition priming.

First of all, we can say that performance on unique stimuli may be expected to follow a power function. We would in the first place expect this from studies of skill acquisition, which have shown that generalized task performance is a power function of practice (e.g., Anderson, 1982; Newell & Rosenbloom, 1981). This expectation is supported by results from the study by Kirsner and Speelman (1996), who examined lexical decision over several blocks of practice, where each block included both novel (i.e., unique) and repeating word–nonword stimuli. They found that power functions provided an excellent fit to performance on unique words and nonwords (Kirsner & Speelman, 1996, Figure 5, and Table 3). Based on these various results, we may conclude that performance on unique stimuli can in general be expected to follow a power function.

As a corollary to the first point, it is clear that, if performance on unique stimuli follows a power function, then skill learning will follow an inverse power function. This follows from the definition of skill learning as being the improvement in performance on unique stimuli over blocks.

What can we say about performance on repeating stimuli? Data pertinent to this question come, once again, from Kirsner and Speelman (1996), who showed that power functions also provided an excellent fit to performance on the repeating words and nonwords in their lexical decision task (Kirsner & Speelman, 1996, Figure 5, and Table 3). Further evidence that multiple repetitions give rise to a reaction time function that follows the power law comes from Logan (1990, Table 1, and Figures 5 and 6). Although this result has sometimes been interpreted to mean that repetition priming follows a power function, this is not strictly true. The specific measure to which Logan's (1990) finding of a power function applies is what we call performance on repeating stimuli (see the earlier discussion of measures used by various authors). Based on the evidence from both Logan and Kirsner and Speelman, then, we may conclude that performance on repeating stimuli can be expected to follow a power function.

As a corollary to the expected power function for performance on repeating stimuli, benefit of repetition is likely to follow an inverse power function.

This leaves the repetition priming measure. First of all, it is worth reiterating that, contrary to some interpretations (e.g., Kirsner & Speelman, 1996), the data in Logan (1990) do not provide evidence about whether the repetition priming measure follows a power function; what they show, rather, is a power-function speedup for performance on repeating stimuli. Let us therefore consider what form the repetition priming function can be expected to take. We begin by noting that repetition priming is the difference between performance on unique stimuli and performance on repeating stimuli, which means that it is the difference between two power functions. The question of what form we can expect for repetition priming thus translates into the question of what form the difference between two power functions can be expected to take. Is there any particular shape for such a function?

The general form of a power function is given by

$$RT = a + bP^{c}, \tag{1}$$

where a is the asymptote, b is the difference between initial performance and the asymptote, P represents the amount of practice, and c represents the rate at which performance improves with practice, $-1 < c < 0$.

Consider two power functions:

$$RT_{\text{unique}} = a_u + b_u P^{c_u}$$

and

$$RT_{\text{repeat}} = a_r + b_r P^{c_r}.$$

Now, the difference between these two functions is repetition priming:

$$\begin{aligned} RT_{\text{unique}} - RT_{\text{repeat}} &= a_u + b_u P^{c_u} - a_r - b_r P^{c_r} \\ &= (a_u - a_r) + P^{c_u}(b_u - b_r P^{c_r - c_u}). \end{aligned}$$

This difference reduces to the general form (1) of a power function iff $c_r = c_u$. Even in this case, for the resultant power function not to be degenerate, we must require that $b_u \neq b_r$, or else the resultant power function reduces to a constant.[2] There is no reason to expect that these conditions will always be satisfied for the power functions that specify performance on unique stimuli and performance on repeating stimuli. Therefore, there is no reason to expect that the difference between these two functions, that is, repetition priming, should follow a power function.

These observations allow resolution of a number of apparently problematic consequences of the recent article by Kirsner and Speelman (1996), to which we have referred several times. The reader will recall that that study examined skill learning and repetition priming in a lexical decision task. Kirsner and Speelman's findings for words are redrawn in Figure 3 and can be summarized as follows. First, performance on unique stimuli shows improvement with experimental practice and follows a power function (Kirsner & Speelman, 1996, p. 568, and Figure 5). Second, performance on repeating stimuli shows improvement with experimental practice and follows a power function (Kirsner & Speelman, 1996, p. 567, and Figure 3). Third, repetition priming does not show an increase beyond the first session and thus does not follow a power function (Kirsner & Speelman, 1996, p. 570, and Figure 7).

The crucial finding was that, unlike skill learning, repetition priming did not follow the power law of practice. Kirsner and Speelman (1996) interpreted this finding as being problematic in two ways. First, they noted that these results appear to directly contradict Logan (1990), who found that repetition priming does follow the power law. Second, Kirsner and Speelman interpreted this finding as strongly challenging the view that skill learning and repetition priming arise from a single mechanism.

With regard to the first point, it should be clear from our earlier discussion that Kirsner and Speelman's (1996) results do not contradict Logan (1990). As we have previously noted, Logan's results demonstrate a power function for performance on repeating stimuli (and for benefit of repetition); they do not demonstrate a power function for the measure that we (and Kirsner and Speelman) term repetition priming. The lack of a power function for this

[2] We thank Gordon Logan for this observation.

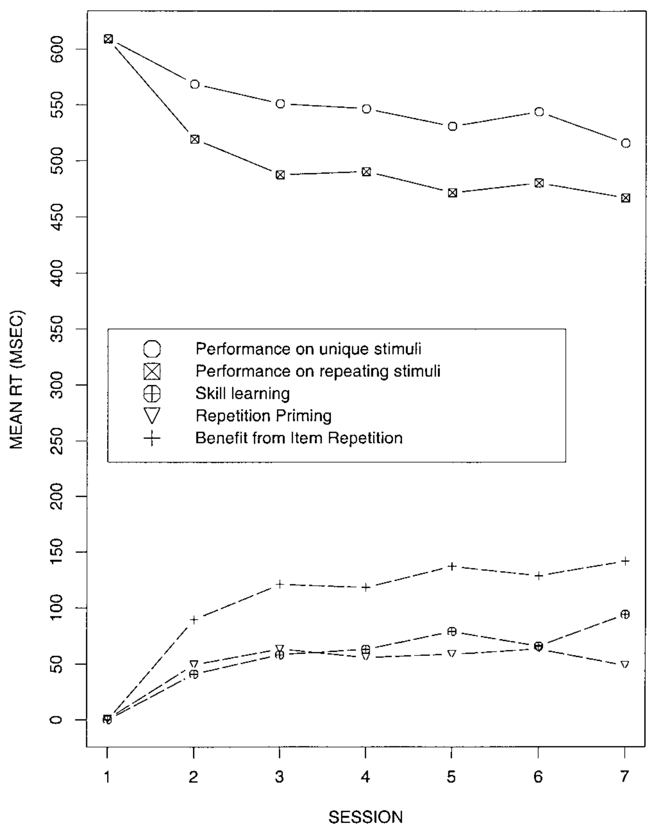


Figure 3. Skill learning and repetition priming results reported for words by Kirsner and Speelman (1996). Data are redrawn from Kirsner and Speelman (1996, Table 2) by collapsing across frequency levels. RT = reaction time.

measure in Kirsner and Speelman's results is therefore in agreement with Logan's findings. Kirsner and Speelman's results also agree with those of Logan in showing a power curve for both performance on repeating stimuli and benefit of repetition. The apparent contradiction between the two experiments arises because Kirsner and Speelman appear to regard the measure for which Logan demonstrated a power function as being the same as the repetition priming measure for which they fall to obtain a power function. However, as we have shown, these measures are not in fact the same. In summary, although one of the main theoretical points made by Kirsner and Speelman was that their data contradicted the results obtained by Logan, we can now see that this claim arises from an unfortunate terminological ambiguity in the literature, whereby different investigators have referred to different measures as "repetition priming." We believe this point underscores the importance of maintaining terminological clarity in discussion of results in this domain.

This brings us to Kirsner and Speelman's (1996) second theoretical argument: that the lack of a power law for repetition priming indicates different underlying mechanisms for skill learning and repetition priming. The reasoning underlying this conclusion is as follows: If repetition priming really were related to skill learning, then, like skill learning, it necessarily would follow a power function. Given their finding that repetition priming does not necessarily follow a power function, Kirsner and Speelman concluded that repetition priming is based on a different principle (and process) than skill learning is.

However, as we have shown above, the fact that the repetition priming measure does not follow the power law follows from its definition and from the fact that performance on repeating stimuli and performance on unique stimuli each follow a power function. This lack of a power law for the repetition priming measure is a definitional artifact and does not bear on the question of whether the mechanisms underlying skill learning and repetition priming are the same.

There appear to have been three reasons for the origin of the erroneous expectation that, if the mechanisms underlying skill learning and repetition priming are indeed the same, then, like skill learning, repetition priming should obey the power law of practice. First, the term repetition priming has not always been used to refer to the same measure. Second, and partly as a result of the first effect, the intuition that the mechanisms that underlie skill learning and repetition priming should be related has been erroneously translated into the expectation that the skill learning and repetition priming measures should be related. Third, the fact that the repetition priming measure is a difference score appears sometimes to have been overlooked.

We turn now to an analysis of similar issues with respect to correlations between skill learning and repetition priming.

Implications of Correlations

As we noted earlier, correlation is a second issue with respect to which the relationship between skill learning and repetition priming has been examined. We will begin by examining data from a digit-entering task. This paradigm was originated by Fendrich, Healy, and Bourne (1991). Here, we will focus on a version of the digit-entering task that was used in our empirical work (Poldrack, Selco, Field, & Cohen, 1999), in which five-digit number strings (e.g., "49385") were presented individually on a computer display to participants, who entered these number strings using a numeric keypad.

Stimulus presentation was divided into blocks. Some digit strings in each block were repeating items. These appeared once in each block and therefore appeared multiple times during the experiment. Some digit strings in each block appeared only once during the experiment (unique items). The structure was therefore identical to that of the hypothetical task that was described at the beginning of this article. Figure 4a shows performance on unique stimuli, performance on repeating stimuli, skill learning, and repetition priming, redrawn from data in the digit-entering task of Poldrack, Selco, Field, and Cohen (1999, Experiment 1). Performance is measured in terms of reaction time.

What Figure 4a shows clearly is that the skill learning and repetition priming measures appear to move in opposite directions across blocks. This pattern is so clear that the two functions could almost be mirror images of each other. Clearly, across a window of several blocks, these two functions would be negatively correlated. Furthermore, this is not an isolated effect that is observed only in the digit-entering task. To illustrate this, we have redrawn data from the three tasks reported as Experiment 1 by Schwartz and Hashtroudi (1991). Figure 4b–d plot these data. Note that in these experiments, performance on unique stimuli and performance on

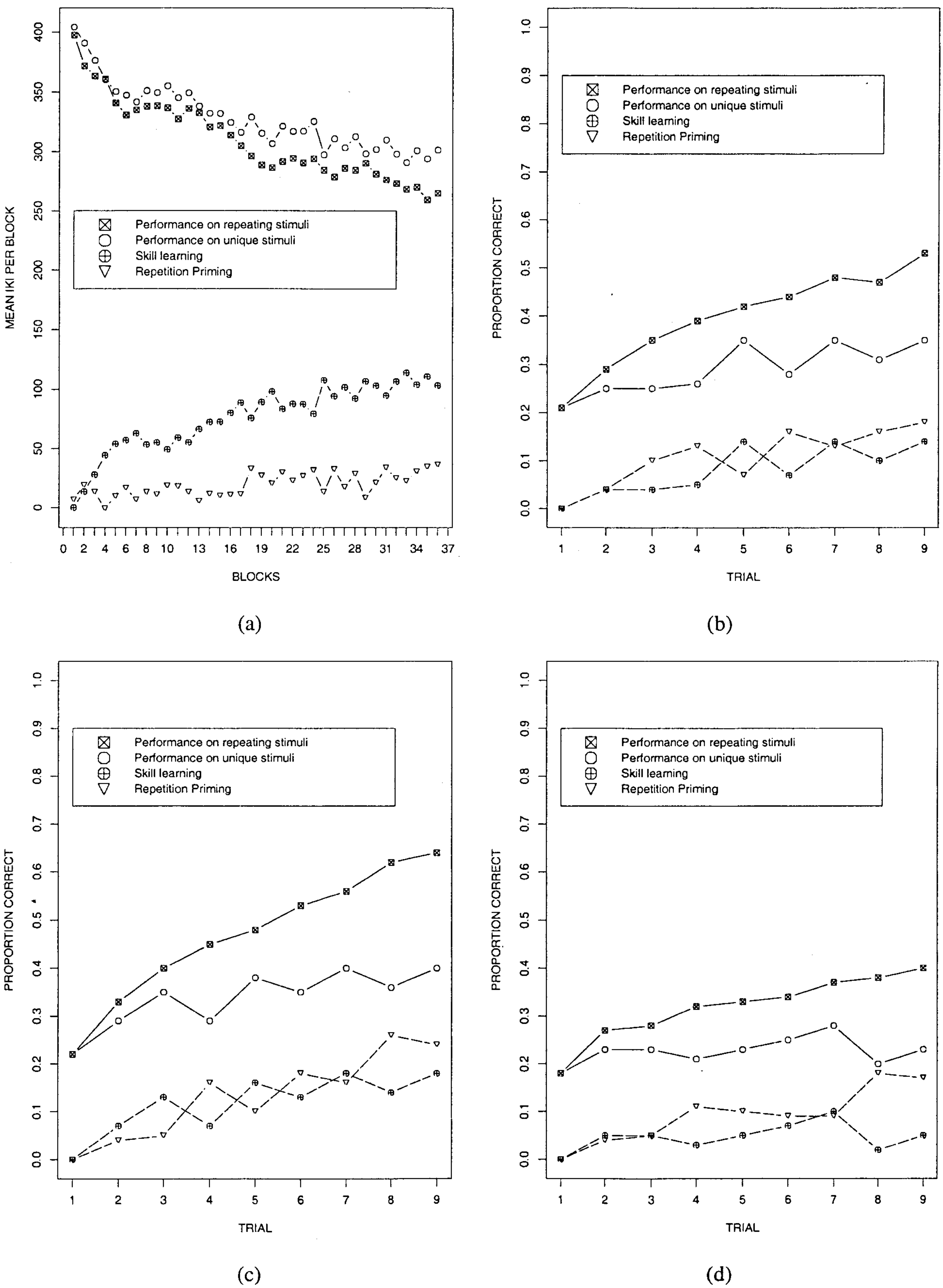


Figure 4. Patterning of changes in skill learning and repetition priming, across blocks. (a) Results from the digit-entering task of Poldrack et al. (1999, Experiment 1). (b) Partial-word identification task. (c) Inverted reading task. (d) Word-fragment completion task. (b)–(d) Results from Schwartz and Hashtroudi (1991, Experiment 1), redrawn. IKI = interkeystroke interval.

repeating stimuli are measured in terms of accuracy rather than in terms of reaction time (see discussion of Figure 2a vs. Figure 2b in *Framework and Terminology*). Accordingly, the performance on unique stimuli and performance on repeating stimuli functions increase with practice rather than decrease with practice, as is the case with reaction time measures. However, the relationships between performance on unique stimuli, performance on repeating stimuli, skill learning, repetition priming, and benefit of repetition are the same in the two situations (see the earlier discussion). Figure 4b–d show that the data from all three of Schwartz and Hashtroudi's (1991, Experiment 1) tasks exhibit the same trend as that observed in Poldrack et al.'s (1999) data: skill learning and repetition priming tend to move in opposite directions. The same trend is also visible in the data of Kirsner and Speelman (1996), shown here in Figure 3. Clearly, the anticorrelation over time of skill learning and repetition priming is quite a general phenomenon.

What is the significance of these observations? In analyzing these patterns of correlation, we need to keep in mind the nature of the various measures involved. Recall that the definitions of skill learning and repetition priming at Block N are $\textit{skill learning}_N = \textit{performance on unique stimuli}_1 - \textit{performance on unique stimuli}_N$ and $\textit{repetition priming}_N = \textit{performance on unique stimuli}_N - \textit{performance on repeating stimuli}_N$. Examining the skill learning measure, we see that skill learning is inversely proportional to performance on unique stimuli. If performance on unique stimuli increases from Block N to Block $N + 1$, skill learning will necessarily decrease. Examining the repetition priming measure, however, we see that repetition priming is directly proportional to performance on unique stimuli. If performance on unique stimuli increases from Block N to Block $N + 1$, repetition priming will necessarily increase also, unless performance on repeating stimuli happens to increase by an even greater amount. As there is no general reason for variation in performance on repeating stimuli to be an amplification of the variation in performance on unique stimuli, repetition priming will usually vary positively with performance on unique stimuli. Thus skill learning varies negatively with performance on unique stimuli, while repetition priming varies positively with performance on unique stimuli. The fact that repetition priming and skill learning tend to move in opposite directions over any window of several blocks is therefore purely a definitional artifact. It has no bearing on whether the mechanisms that underlie skill learning and repetition priming are the same or different. That is, it has no theoretical significance. But this raises questions about the significance of correlations between the skill learning and repetition priming measures at any particular block; after all, a particular block is equivalent to a window size of one. Is there any reason why skill learning and repetition priming should be correlated across participants, on any given block? This question is worth examining in greater detail.

To examine more closely participants' performance in a particular block, we arbitrarily selected Block 6 of the digit-entering task data from Experiment 1 of Poldrack et al. (1999). Block 6 performance on unique stimuli and performance on repeating stimuli is shown for each of the participants in the experiment in the upper panel of Figure 5a. As can be seen, the two measures are highly correlated across participants, for this arbitrarily chosen block ($r = .927$, $p < .001$). The lower panel of Figure 5a repeats the performance on unique stimuli and performance on repeating stimuli curves from the upper panel, additionally showing each participants' performance on unique stimuli in Block 1 of the experiment, this being the baseline measure with respect to which skill learning is computed; recall that skill learning on a particular block is measured as the difference between performance on unique stimuli on that block and performance on unique stimuli in Block 1. This difference, that is, skill learning, is also plotted in the lower panel of Figure 5a. Finally, the lower panel also shows each participant's repetition priming in Block 6, computed as the difference between performance on unique stimuli and performance on repeating stimuli for each participant. Examination of the skill learning and repetition priming functions reveals that the correlation between these measures is likely to be weak or nonexistent, and this is in fact the case ($r = .222$, $p > .4$). This pattern of results also holds for other arbitrarily chosen blocks, such as Block 18, shown in Figure 5b. We can see that in Block 18, as in Block 6, the performance on unique stimuli and performance on repeating stimuli measures are strongly positively correlated across participants (upper panel; $r = .892$, $p < .001$), but that the skill learning and repetition priming measures have a weak or negative correlation (lower panel; $r = -.269$, $p > .3$).

These analyses provide an understanding of the significance of the various correlations by making it clear that both skill learning and repetition priming are derived measures. In any given block, each participant's skill learning is derived by subtracting the participant's performance on unique stimuli in that block from the participant's performance on unique stimuli in Block 1. In consequence, the form of the skill learning function across participants diverges from the form of performance on unique stimuli, even though skill learning is a measure of improvement in performance on unique stimuli. Repetition priming for each participant is derived as the difference between performance on unique stimuli and performance on repeating stimuli for that participant. As a result, the form of the function across participants diverges from that of the performance on repeating stimuli function, even though repetition priming is a measure of improvement in performance on repeating stimuli. Because of these divergences, there is no logical reason why the derived measures should be correlated. It is worth noting that these results are quite general. They are not restricted to data from the particular blocks shown in Figure 5a and 5b. Nor are they even restricted to the digit-entering task. These are functional arguments that apply to skill learning and repetition priming in any task in which they are jointly observed. There is no logical reason why skill learning and repetition priming should be correlated in any given block or in any given task.

This can be confirmed by examining the patterns of correlation systematically at every block. Figure 6 shows correlations across participants at each block in the digit-entering task. Note that the correlations for Blocks 6 and 18 are the same as those obtained from Figure 5a and 5b. The upper line shows that correlation between performance on unique stimuli and performance on repeating stimuli is highly positive at all blocks in the experiment; we will return to this point later. More important for the present discussion, the lower line shows that the skill learning and repetition priming measures are negatively correlated for the most part, as our earlier analysis would suggest. We therefore would find no positive correlation in most of the blocks of the digit-entering task or in the partial-word identification task or the inverted reading task of Schwartz and Hashtroudi (1991, Experiment 1). Neverthe-

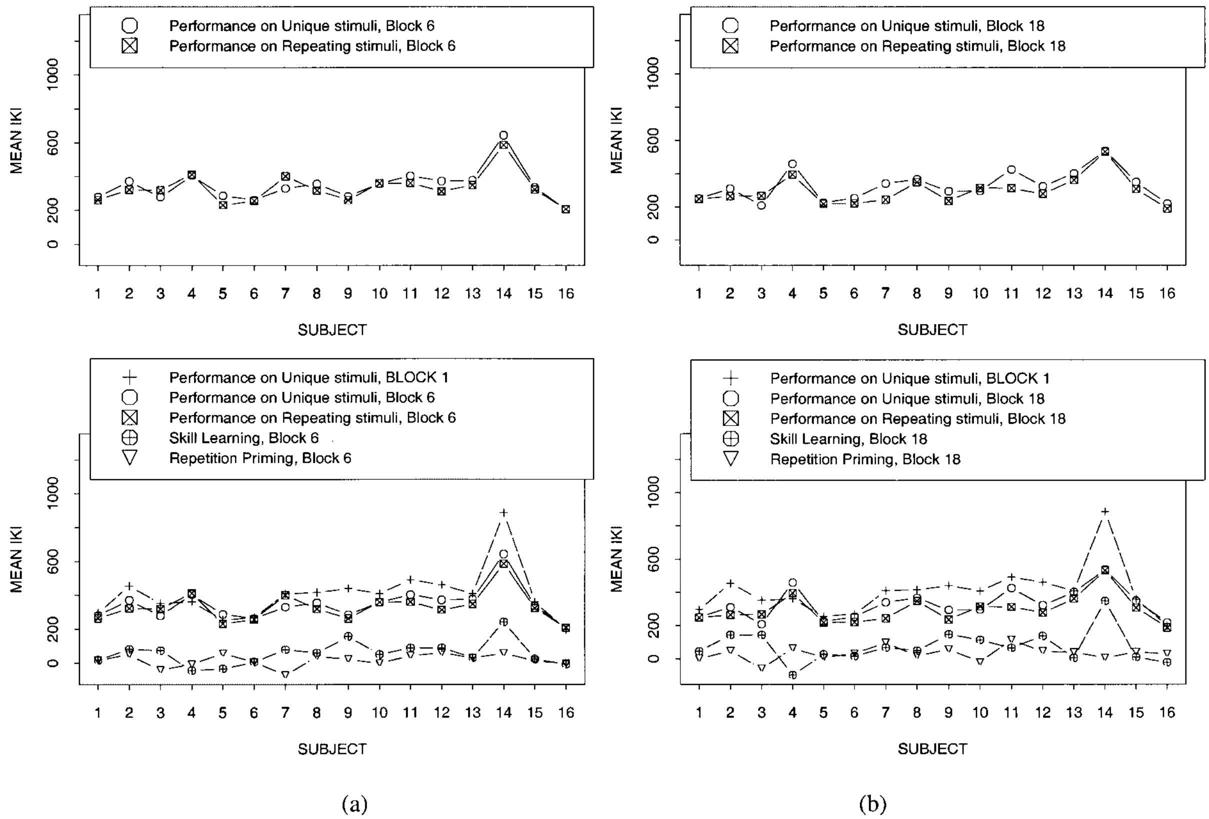


Figure 5. Analysis of correlations from the digit-entering task. (a) Participant-wise results at Block 6. (b) Participant-wise results at Block 18. IKI = interkeystroke interval.

less, these two measures can on occasion be positively correlated in a particular block, as is true in Blocks 34 and 36 of the digit entering task (Figure 6), and in the word-fragment completion task of Schwartz and Hashtroudi (1991, Experiment 1), where a positive correlation between skill learning and repetition priming was obtained in the last block of the experiment. However, a positive or negative correlation or lack of correlation between these two measures on any given block has no theoretical significance.

We believe that interpretation of the lack of correlation between the skill learning and repetition priming measures as being indicative of a processing dissociation does not take into consideration the nature of the measures. The analysis in this section should make it clear that the presence, absence, or strength of correlations between the skill learning and repetition priming measures is artifactual and follows from their definitions. It has no theoretical significance. Therefore, the discovery of a lack of correlation between the skill learning and repetition priming measures in some tasks but not in others (Schwartz & Hashtroudi, 1991) does not constitute evidence that the mechanisms underlying skill learning and repetition priming are necessarily different.

Relevant Measures

Our discussion so far has shown that patterns of dissociation between the skill learning and repetition priming measures do not warrant conclusions about the underlying mechanism or mechanisms. This leaves the question of how the relation (or lack thereof) between the unknown mechanisms may be examined. We suggest that, in thinking about such processing relationships, one needs to take into account not just the standard measures of skill learning and repetition priming, but also the measures we have termed performance on unique stimuli and performance on repeating stimuli.

There are at least two important reasons why performance on unique stimuli and performance on repeating stimuli should be taken into consideration. First, it is clearly a good idea to take into account a larger rather than a smaller number of measures. Our analyses in the preceding sections clearly demonstrate the importance of this approach: As we showed, it is only by taking into account the performance on unique stimuli and performance on repeating stimuli measures that it becomes clear why patterns of dissociation between the standard measures of skill learning and repetition priming are uninformative about the underlying processing.

A second reason is that the performance on unique stimuli and performance on repeating stimuli measures together encode more information than do the standard skill learning and repetition priming measures. If we are given only the performance on unique stimuli and performance on repeating stimuli functions that are obtained in a task, we can automatically derive the skill learning and repetition priming functions. The converse is not true, however: The performance on unique stimuli and performance on repeating stimuli functions cannot be derived given only the skill

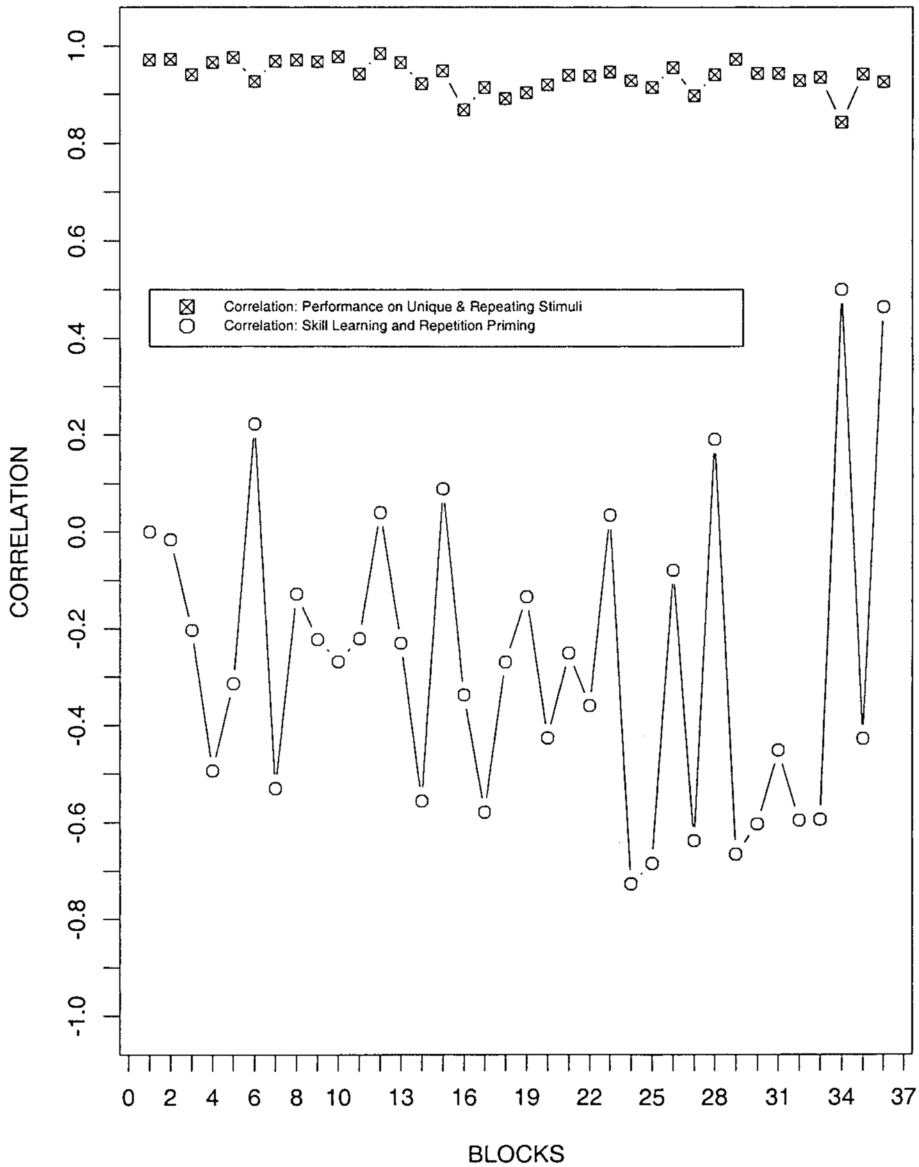


Figure 6. Analysis of correlations at each block in the digit-entering task.

learning and repetition priming functions for a task. The significance of this becomes clearer when we consider the following: A processing theory that accounted for the performance on unique stimuli and performance on repeating stimuli functions in a task would necessarily account for the skill learning and repetition priming functions in that task. However, it is not clear whether a processing theory that accounted for the skill learning and repetition priming functions in a task would necessarily account for the actual performance data (i.e., performance on unique stimuli and performance on repeating stimuli). The performance on unique stimuli and performance on repeating stimuli curves thus constitute the primary data that must be explained when constructing a processing account. It therefore seems highly desirable to take these measures into account when making processing inferences from patterns of data that are obtained in a behavioral manipulation.

These considerations make a strong argument for taking the performance on unique stimuli and performance on repeating stimuli measures into account in investigations of skill learning and repetition priming. It is important to note also that these arguments are theory independent. They retain their force irrespective of the validity of any particular theory of skill learning and repetition priming. In addition to these theory-neutral arguments, there are also some theory-dependent reasons to believe in the importance of taking the performance on unique stimuli and performance on repeating stimuli measures into consideration. These reasons follow from the theory of procedural memory that we outline in the second part of this article which, however, requires some discussion here.

In our theory, procedural learning in a particular task is viewed as occurring in whatever processors are required for performance of that task. These processors can be thought of loosely as the "neural assemblies" that need to be deployed in performance of that particular task. Clearly, the specific processors that are deployed will differ widely across different tasks. The particular set of processors that are required for performance of a particular task collectively constitute the "system" that performs that task. Important in our view is that there is no difference in how the system that is necessary for a particular task operates for different types of stimuli that are processed in that same task (such as repeating vs. unique stimuli). All of these stimuli are processed by exactly the same system (i.e., the specific set of processors that are needed for that task). Furthermore, each time any of these stimuli is processed by the system, there is incremental "tuning" of the processors, which adjusts the system in such a way as to facilitate subsequent processing of that stimulus by the system. Note that this incremental tuning occurs each time a stimulus is processed and occurs in the same way irrespective of the type of stimulus. A good analogy for this tuning is the kind of connection weight adjustment that occurs in a connectionist network (e.g., of the kind described in Rumelhart, Hinton, & Williams, 1986). In such networks, adjustment of connection weights following presentation of a particular stimulus facilitates the network's performance on a subsequent presentation of that stimulus.

This tuning process leads not only to facilitated processing on the specific stimuli to which the system is exposed, but also to facilitated processing on stimuli from the given membership class in general (provided there is some basis for the representation of similarity between stimuli, such as that provided by distributed representations, which is discussed in the *Description of the model* subsection). Thus, the continual tuning process adjusts the system to do better, in general, on the task. But it adjusts the system even more for those stimuli that repeat, precisely because they repeat: The system undergoes greater attunement favoring these stimuli than it does for stimuli in general. The underlying system, the learning mechanism, and the state of tuning of the system are nevertheless the same for the two kinds of stimuli. In terms of the analogy with the connectionist network, all stimuli (be they "unique" or "repeating") are processed by exactly the same units, connections, and connection weights. If we observe the system's performance on unique versus repeating stimuli, we will usually see that there is generalized facilitation on the unique stimuli and even greater facilitation on the repeating stimuli. The output of the system will, in other words, generate performance on unique stimuli and performance on repeating stimuli functions. This theoretical view is made rigorously operational in the form of a computational model and simulations in the second part of this article. For now, what we wish to emphasize is that, in our theory, unique and repeating stimuli are not processed differently. There is neither a separate learning mechanism nor separate sets of processors for these (or any other) different types of stimuli. What then is the difference between repeating and unique stimuli? In our view, the only difference is the degree of facilitation the system undergoes in processing these types of stimuli. And this degree of facilitation is determined by how many times the system has been exposed to each type of stimulus.

Skill learning is in our view merely the degree of facilitation on the unique items, and repetition priming is merely the additional degree of facilitation that arises from the additional exposure of the system to the repeating items. Both kinds of facilitation are encoded by a single state of "tuning" in the system (in the neural network analogy, by a single set of connection weights) that determines processing of stimuli in the system. Underlying both kinds of facilitation is the incremental adjustment mechanism that is always operative throughout the system and that is blind to whether a particular stimulus is unique or repeating. We term this mechanism *procedural learning*; the changes arising from procedural learning are observed as *procedural memory* effects.

It should be clear why, in this view, the performance on unique stimuli and performance on repeating stimuli functions are considered important. They directly reflect the the current state of tuning of the system (which directly determines the operation and performance of the system). The standard measure of repetition priming does not directly reflect the current tuning of the system; rather, it reflects the difference between how well the system is tuned for (and therefore performs on) stimuli from the membership class in general versus those stimuli that have repeated. The standard measure of skill learning does not directly reflect the current state of tuning (and therefore processing) in the system; rather, it reflects the extent to which the current tuning of the system has changed with respect to some earlier point in time.

On this view, we would expect the performance on unique stimuli and performance on repeating stimuli measures to exhibit similar patterns in a given task, and in fact this appears to be the case, as we have already seen. The two measures exhibit the same general trend in the data of Schwartz and Hashtroudi (1991; see Figure 2b, this article); in Kirsner and Speelman's (1996) data (see Figure 3, this article); and in the data of Poldrack et al. (1999, Experiment 1; see Figure 4a, this article). They also exhibit a strikingly close relationship across participants at two arbitrarily chosen blocks in the data of Poldrack et al. (1999, Experiment 1), as shown graphically in Figure 5a and 5b. This indication was confirmed in Figure 6, which shows that the two measures have a correlation of close to 1.0 at each block. Furthermore, these two measures both follow power functions, as discussed earlier.

This does not mean that the standard measures of skill learning and repetition priming are irrelevant or less important than are the meaures of performance on unique stimuli and performance on repeating stimuli simply because they are derived or baseline adjusted. Indeed, to measure skill learning (defined as the extent of facilitation in generalized task performance), the baseline must be subtracted out, so the standard measure of skill learning is the appropriate one. To measure repetition priming (defined as the extent of facilitation resulting from repetition of specific stimuli, over and above facilitation due to generalized improvement), skill learning must be subtracted out, so that the standard measure of repetition priming is the appropriate one. However, although these measures are perfectly valid, patterns of dissociation between them are less informative about the underlying processing than is rather widely supposed, as our theory-neutral analysis in the preceding two sections clearly shows.

In summary, performance on unique stimuli and performance on repeating stimuli are important to keep in mind even on completely theory-independent grounds. In addition, according to our particular theoretical framework, there are further reasons to regard performance on unique stimuli and performance on repeating stimuli as particularly relevant measures. With these points in mind, we now turn to examining other patterns of dissociation that have been reported in the implicit memory literature.

Patterns of Increase and Decrease

A third kind of relationship between skill learning and repetition priming that has been examined is the relationship between their patterns of change. As we noted earlier, this question amounts to asking whether skill learning and repetition priming are correlated over the course of an entire experiment (rather than across participants at a particular point in the experiment). The underlying intuition here has been that, if skill learning and repetition priming are related, then they should both increase or decrease together, over the course of experimental practice (e.g., Schwartz & Hashtroudi, 1991). What is the theoretical significance, then, of a pattern of relationship in which one measure increases with practice, but the other does not?

We can begin by considering possible relationships between performance on unique stimuli and performance on repeating stimuli. For each of these measures, there are two possibilities with respect to the pattern of change with practice: performance either (a) improves or (b) stays constant. We will ignore the possibility that performance deteriorates with practice. Such deterioration might in fact occur as a result of fatigue; however, we regard fatigue effects as extrinsic. That is, the facilitation of processing is an intrinsic effect of practice, whereas fatigue is an extrinsic effect of task duration. Fatigue might temporarily mask facilitation, but the facilitation would be revealed if the task were resumed after a rest period. This can be thought of as a competence-performance distinction, and for our theoretical analysis we consider only the underlying competence and its relatively long-lasting facilitation by practice or repetition. We also ignore deterioration in performance that arises from a qualitative change in the way a task is performed (such as a strategy shift) or from a change in the representational status of stimuli. An example of the latter is the decrease in lexical decision accuracy for repeating nonwords reported by Kirsner and Speelman (1996), which in our view arose because the repeating nonwords attained the representational status of "words," which interfered with their correct rejection. Thus, confining our analysis to the possibility that performance either improves or stays constant, there are four possible relationships between performance on unique stimuli and performance on repeating stimuli: (a) Both performance on unique stimuli and performance on repeating stimuli improve with practice; (b) Performance on unique stimuli is constant, while performance on repeating stimuli improves; (c) Performance on unique stimuli and performance on repeating stimuli are both constant; and (d) Performance on unique stimuli improves, while performance on repeating stimuli is constant. We will assume that performance on unique stimuli and performance on repeating stimuli are measured in terms of reaction time, so that an improvement in performance translates into a decrease in reaction time, with practice. The four possible relationships between performance on unique stimuli and performance on repeating stimuli (measured by reaction time) are represented by the four cells in Figure 7. Within these cells, there are some further possibilities. Each possibility is shown in terms of its implications for the performance on unique stimuli and performance on repeating stimuli curves as well as for the derived skill learning and repetition priming curves.

Within Case I (performance on unique stimuli improving, performance on repeating stimuli improving), there are four logical possibilities, as shown in the figure: (Ia) The performance on unique stimuli and performance on repeating stimuli functions are parallel. This implies increasing skill learning, with constant repetition priming. (Ib) The performance on unique stimuli and performance on repeating stimuli functions diverge. This implies that both skill learning and repetition priming increase. (Ic) The performance on unique stimuli and performance on repeating stimuli functions converge. This implies that skill learning increases, while repetition priming decreases. (Id) The performance on unique stimuli and performance on repeating stimuli functions are identical. This implies that skill learning increases but repetition priming is zero. Case II has one logical possibility: Performance on unique stimuli is constant, and performance on repeating stimuli improves, so that skill learning is zero, while repetition priming increases. Within Cell III (performance on unique stimuli constant, performance on repeating stimuli constant), there are two possibilities: (IIIa) The performance on unique stimuli and performance on repeating stimuli functions are parallel. This implies that skill learning is zero, and repetition priming is constant. (IIIb) The performance on unique stimuli and performance on repeating stimuli functions are identical. This implies that skill learning and repetition priming are both zero. Finally, in Case IV, performance on unique stimuli improves while performance on repeating stimuli is constant, so that skill learning increases while repetition priming decreases.

The eight patterns identified above constitute all logically possible patterns of relationship between performance on unique stimuli and performance on repeating stimuli. What, if anything, does each of these possible patterns signify about the mechanisms underlying skill learning and repetition priming? Cases Ia–Id exhibit a variety of patterns of dissociation between the skill learning and repetition priming measures. In Case Ia, for instance, skill learning increases whereas repetition priming does not—the pattern reported by Kirsner and Speelman (1996). In Case Ic, skill learning increases whereas repetition priming actually decreases. In Case Id, skill learning increases while repetition priming is zero. Such patterns of dissociation might be thought to indicate separate underlying mechanisms.

However, when we examine the performance on unique stimuli and performance on repeating stimuli functions in Cases Ia–Id, there appears to be a close relationship between these measures. In each case, performance on unique stimuli and performance on repeating stimuli both have the form of power functions. (Note, however, that the analyses below would hold even if the performance on unique stimuli and performance on repeating stimuli functions were not power functions, for instance, if they were linear.) Furthermore, the performance on unique stimuli and performance on repeating stimuli measures are highly correlated in each case. The same can be said of Cases IIIa and IIIb: In each case, the performance on unique stimuli and performance on repeating stimuli functions have the same form—they are constant and can be thought of as degenerate power functions. They are also highly correlated. Thus, in contrast to the dissociations exhibited by the skill learning and repetition priming measures, the performance on unique stimuli and performance on repeating stimuli measures exhibit similar patterns to each other in the various scenarios in Cases I and III.

What implications do these various patterns of relationship have for the underlying processing? Are we to draw inferences from the patterns of dissociation between the skill learning and repetition

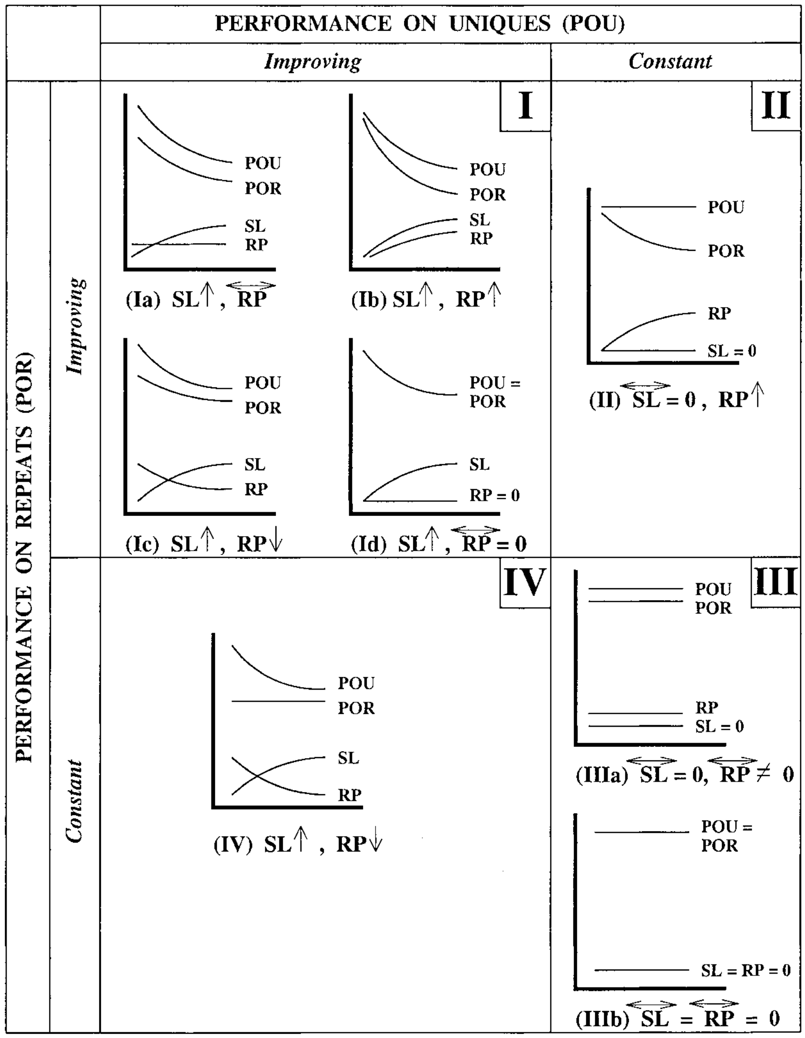


Figure 7. Possible relationships between measures of performance on unique stimuli (POU), performance on repeating stimuli (POR), skill learning (SL), and repetition priming (RP).

priming measures or from the patterns of similarity between the performance on unique stimuli and performance on repeating stimuli measures? In the previous section, we pointed out the importance of the performance on unique stimuli and performance on repeating stimuli measures, even on purely theory-independent grounds: We noted that any processing account would have to account for the performance on unique stimuli and performance on repeating stimuli data and in doing so would necessarily account for the skill learning and repetition priming functions. Additionally, much of our discussion to this point has been concerned with showing that other commonly reported dissociations between the standard measures of skill learning and repetition priming are artifactual and do not justify inferences regarding the underlying processing. At the very least, therefore, the patterns of similarity between the performance on unique stimuli and performance on repeating stimuli measures in Cases Ia–Id and Cases IIIa and IIIb should give us pause for thought; they need to be explained before we draw the conclusion that the various measures in each scenario arise from different underlying mechanisms.

A stronger interpretation of the patterns of similarity between the performance on unique stimuli and performance on repeating stimuli measures would be that they reflect the operation of a single underlying mechanism. This is, of course, the position that we adopted in our theory of procedural memory, discussed in the preceding section. Even though there is no theory-independent way of establishing this interpretation, it is worth pointing out that our theoretical framework does offer a very specific account of how a single underlying mechanism can give rise to such performance on unique stimuli and performance on repeating stimuli functions. In fact, as we will demonstrate in the second part of the article, our theory also offers an account of why the patterns of dissociation between the skill learning and repetition priming measures would arise even with a single underlying mechanism.

We suggest, therefore, that none of the scenarios in Cases I and III constitutes conclusive evidence that there are separate mechanisms underlying skill learning and repetition priming. What about Cases II and IV? In Case II, performance on unique stimuli is constant while performance on repeating stimuli improves; in Case

Table 1
Possible Patterns of Increase–Decrease in the Skill-Learning (SL) and Repetition-Priming (RP) Measures, Exemplified by the Scenarios Depicted in Figure 7

RP	SL: Increasing	SL: Constant	SL: Decreasing
Increasing	Ib	II	Impossible
Constant	Ia, Id	IIIa, IIIb	Impossible
Decreasing	Ic, IV	Impossible	Impossible

Note. Certain cells denote patterns of relationship that are impossible on our assumptions.

IV, performance on repeating stimuli is constant while performance on unique stimuli improves. In these cases, it is not clear whether the performance on unique stimuli and performance on repeating stimuli measures would be correlated or would have functions of the same form. It might be argued that empirical observation of these scenarios would constitute evidence that skill learning and repetition priming arise from different mechanisms, because even the performance on unique stimuli and performance on repeating stimuli measures appear to be dissociated.

In a later section, we demonstrate that the patterns of relationship in Cases II and IV can in fact arise from a single mechanism. We describe a simulation in which repetition priming increases in the absence of skill learning in the same computational model, which corresponds to the pattern of results in Case II. This will demonstrate that the seemingly disparate skill learning and repetition priming functions in Case II do not necessarily imply different mechanisms. We also describe a simulation in which performance on unique stimuli improves with practice, whereas performance on repeating stimuli exhibits virtually no improvement, corresponding to the pattern of results in Case IV, but which arises in a single model. This will show that the seemingly very different performance on unique stimuli and performance on repeating stimuli functions shown in Case IV do not necessarily imply different mechanisms either. Further elaboration of these simulations requires detailed description of the computational model. We felt that incorporating this description at the present point in the discussion would be too digressive. The reader is therefore asked to accept on faith our assertion that the relationships between performance on unique stimuli and performance on repeating stimuli that are depicted in Cases II and IV do not necessarily imply the existence of different underlying mechanisms. This promissory note will be made good in the *Simulated Patterns of Increase and Decrease* subsection.

What we can say, then (if our promissory note is accepted) is that none of the possible patterns of relationship between performance on unique stimuli and performance on repeating stimuli or between skill learning and repetition priming as depicted in Cases I–IV is necessarily inconsistent with the possibility that a single mechanism underlies skill learning and repetition priming. Notice further that Cases I–IV include every possible pattern of relationship between the skill learning and repetition priming measures, as shown in Table 1. In Case Ib, skill learning and repetition priming both increase. In Cases Ia and Id, skill learning increases while repetition priming is constant. In Cases Ic and IV, skill learning increases while repetition priming decreases. In Case II, skill learning is constant while repetition priming increases. In Cases IIIa and IIIb, skill learning and repetition priming are both constant. The remaining cells shown in Table 1 are ruled out by our assumption that performance on unique stimuli and performance on repeating stimuli cannot deteriorate with practice.[3] Thus, all possible patterns of relationship between the skill learning and repetition priming measures are covered by the scenarios in Cases I–IV. We have argued that none of the scenarios in Cases I and III constitutes conclusive evidence that there are different mechanisms that underlie skill learning and repetition priming. We have promised to show also that neither of the scenarios in Cases II and IV is inconsistent with the possibility that the phenomena of skill learning and repetition priming arise from a single mechanism.

The import of these observations is that none of the possible patterns of relationship between the standard measures of skill learning and repetition priming is necessarily inconsistent with the possibility of a single underlying mechanism. Stated differently, no matter which of the relationships between the skill learning and repetition priming measures is obtained in an experiment, it does not unequivocally establish that the phenomena of skill learning and repetition priming are based on different mechanisms.

Discussion

In the preceding sections, we developed a theoretical analysis of various kinds of relationship between the standard measures of skill learning and repetition priming. We provided an analysis of their relationship in terms of the shape of their functions, their correlation, and the patterning of their rates of change. These analyses highlight the importance of taking into consideration the measures we term performance on unique stimuli and performance on repeating stimuli and keeping in mind how the standard skill learning and repetition priming measures are defined. When the nature of the performance on unique stimuli, performance on repeating stimuli, skill learning, and repetition priming measures is taken into account, it becomes clear that patterns of relationship between the skill learning and repetition priming measures do not have any theoretical significance. In particular, a lack of the expected relationships does not constitute evidence for separate mechanisms. In our analyses, we demonstrated this in several ways. First, we showed that the fact that the repetition priming measure does not follow the power law follows from its definition and from the fact that performance on repeating stimuli and performance on unique stimuli each follow a power function. The lack of a power law for the repetition priming measure is a definitional artifact; there is no reason to expect this measure to exhibit any specific form. Second, we showed that the presence or absence of correlations between the skill learning and repetition priming measures is artifactual and follows from their very definitions. Third, we argued, with respect to patterns of increase-

[3] To see this, note that all the cells in the rightmost column of Table 1 involve decreasing skill learning, which implies that performance on unique stimuli deteriorates with practice. The cell corresponding to constant skill learning and decreasing repetition priming implies a flat performance on unique stimuli function together with an upward sloping performance on repeating stimuli function; that is, it implies deteriorating performance on repeating stimuli.

decrease, that none of the possible patterns of relationship between the skill learning and repetition priming measures conclusively establishes that there are different underlying mechanisms, and that none of the patterns is necessarily inconsistent with the existence of a single underlying mechanism.

In light of these various analyses, it is possible to reexamine certain findings that have been presented as evidence that skill learning and repetition priming are unrelated. For example, in the results reported by Kirsner and Speelman (1996), the skill learning measure increased with practice, whereas the repetition priming measure was constant (beyond the second session of practice; see Figure 3). The authors interpreted this pattern of relationship as indicating different underlying mechanisms. However, when we view this pattern of results in terms of our framework of possible relationships between performance on unique stimuli and performance on repeating stimuli (Figure 7), we can see that it corresponds to Case Ia. As we have just seen, neither Case Ia nor any of the other possible patterns of relationship between the skill learning and repetition priming measures is necessarily inconsistent with a relatedness of underlying mechanisms.

As already noted, Schwartz and Hashtroudi (1991, Experiment 1) examined repetition priming and skill learning in each of three tasks: partial-word identification, inverted reading, and word-fragment completion. They found that priming effects were not correlated with improvements in skill across trials, in either the partial-word identification or inverted reading tasks, and took these results as indication of separate mechanisms underlying skill learning and repetition priming. In light of the analyses we have presented here, it should be clear that these correlational data have no bearing on the question of whether or not a single mechanism underlies the phenomena of skill learning and repetition priming.

Schwartz and Hashtroudi (1991) also observed as a separate point that, whereas the magnitude of repetition priming was similar across the three tasks, skill learning occurred in only two of the tasks (see Figure 4b–4d). Schwartz and Hashtroudi took this finding also to indicate that skill learning and repetition priming are unrelated. The patterns that were obtained in these three tasks can be classified as Cases Ib, Ib, and II, respectively, in terms of our analyses from Figure 7.[4] As we have argued, these patterns of relationship between the skill learning and repetition priming measures are consistent with the existence of a single underlying mechanism (as indeed are all the possible patterns of relationship). This aspect of Schwartz and Hashtroudi's results therefore does not offer conclusive evidence regarding the issue of single versus dual mechanisms, either.

Schwartz and Hashtroudi (1991, Experiment 2) further examined whether the amount of skill learning in the partial-word identification task would influence the magnitude of priming in that task. They found that the magnitude of priming did not differ significantly for two groups of participants with different levels of previous practice in the task and concluded that priming in partial-word identification is unrelated to the amount of skill in identifying degraded words. Schwartz and Hashtroudi also examined whether word frequency would modulate priming and skill learning. They found that skill learning was significantly greater for the high-frequency words than for the low-frequency words, while the magnitude of repetition priming was unaffected by frequency (Experiment 3). They suggested that these differential effects of word frequency indicate that skill learning and repetition priming are unrelated (p. 1183). Based on the analytical framework that we have established in preceding sections, it can be shown that neither of these interpretations is valid. A detailed demonstration of this is presented in the Appendix. Briefly, the former interpretation is based on an inaccurate task analysis. The latter result is a dissociation that arises from a focus on the skill learning and repetition priming measures; when the performance on unique stimuli and performance on repeating stimuli measures are taken into consideration, it becomes clear that word frequency has almost identical effects on the processing of unique and repeating stimuli.

The utility of our theoretical framework is not confined to the studies that we have discussed so far; rather, it has applicability to interpretation of a wide variety of results regarding skill learning and repetition priming. One example is that it offers a quite general analysis of the relationship between study–test paradigms and multiple-repetition priming paradigms (see discussion in the Appendix) as well as insight into how priming effects may differ across the two kinds of paradigms (see the *Multiple Levels of Skill Learning and Repetition Priming* subsection). Another example that we have already touched on is worth mentioning once again and comes from the work of Logan (1990), who provided evidence that multiple repetitions of stimuli in a lexical decision task give rise to a reaction time function that follows the power law (Logan, 1990, Table 1, and Figures 5 and 6). Logan fitted power functions to the mean reaction time data (Logan, 1990, Table 1, and Figures 5 and 6), noting that "the power function fits reveal an important empirical parallel between automaticity and repetition priming: The learning function that characterizes automaticity also provides a good description of repetition priming" (p. 25). Although these results appear to indicate that repetition priming follows the power law (and this is how Kirsner and Speelman, 1996 interpreted them), our analytic framework makes it easy to see that the measure that is characterized by a power function in Logan's data is what we have termed performance on repeating stimuli. Once this is seen, it becomes clear that the lack of a power function for the standard measure of repetition priming in other studies does not contradict Logan's results; indeed, other studies confirm the existence of a power function for performance on repeating stimuli (e.g., Kirsner & Speelman). This insight is a direct result of applying our analytic framework.

As another example of the applicability of our framework, we can consider results reported by McAndrews and Moscovitch (1990, Experiment 4), who examined participants' anagram-solving ability. The authors found that both the application of a

[4] Note that the performance on unique stimuli and performance on repeating stimuli curves for Schwartz and Hashtroudi's (1991) data in Figures 4b–4d are measured in terms of accuracy, rather than in terms of reaction time. Consequently, the shape of the performance on unique stimuli and performance on repeating stimuli curves is inverted with respect to the depictions of possible scenarios in Figure 7. It should be possible to see, nevertheless, that Figures 4b and 4c correspond to Case Ib in Figure 7: There is a pattern of improving and diverging performance on unique stimuli and performance on repeating stimuli. As a result, the skill learning and repetition priming measures both increase. Similarly, Figure 4d corresponds to Case II in Figure 7: There is a pattern of constant performance on unique stimuli, with improving performance on repeating stimuli. As a result, the skill learning measure is effectively constant at zero, while repetition priming increases.

previously learned implicit rule (i.e., the learning of a skill) and previous experience with a specific anagram (i.e., the effect of repetition) facilitated performance, and the effects of these two sources of facilitation were additive and independent. They concluded that these findings did not support the view that skill learning and repetition priming are supported by the same procedural memory system (p. 784). We believe that our analytical framework makes it possible to see that this statistical independence is not very informative as to the underlying processing. Briefly, our framework makes it clear that the four data points that McAndrews and Moscovitch studied are single points on a family of four performance functions, and that inferences about the mechanisms underlying these performance functions cannot be made from comparisons between single points on these functions. Our analysis is presented in detail in the Appendix.

Additionally, our framework helps shed light on a number of apparent dissociations between skill learning and repetition priming that have been reported in the neuropsychological literature. In the *Relevant Measures* subsection, we outlined our view that the particular set of processors that is required for performance of a particular task collectively constitute the "system" that performs that task, and that the specific processors deployed will in general differ across different tasks. Practice in a particular task can lead to facilitation (which we called "tuning") in any subset of the set of processors subserving that task, and it is this facilitation that is manifested as skill learning or repetition priming. If we compare performance across two tasks, we are comparing facilitation in what are very likely different sets of processors. It therefore may not be appropriate to make inferences about the mechanisms that underlie skill learning and repetition priming that are based on comparisons across different tasks. The importance of making a distinction between tasks and processes and of considering in what way different tasks rely on different combinations of processes has in recent years been emphasized by a number of other investigators as well (e.g., Jacoby & Kelley, 1992; Kirsner, 1998; Roediger, Buckner, & McDermott, 1999; Witherspoon & Moscovitch, 1989). Other investigators have also pointed out that it is critical to examine skill learning and repetition priming within the same task (Schwartz & Hashtroudi, 1991), and, indeed, such considerations formed a primary motivation for the examination of skill learning and repetition priming within a single task in our empirical work on the digit-entering task (Poldrack et al., 1999).

The relevance of this to interpreting neuropsychological dissociations between skill learning and repetition priming is that these have typically involved comparisons across different tasks. For example, Heindel et al. (1989) found that patients with Alzheimer's disease exhibited normal motor skill learning in a pursuit-rotor task, but impaired priming in word- and picture-completion tasks. In contrast, patients with Huntington's disease showed impaired skill learning in the pursuit-rotor task, but normal priming in word- and picture-completion tasks. The authors interpreted these results as showing a dissociation between skill learning and repetition priming (Heindel et al., 1989). An important point that emerges from our framework, however, is that the processors underlying these various tasks are quite different. Therefore (as Schwartz & Hashtroudi, 1991, have also pointed out), a dissociation between word completion and pursuit-rotor tasks may occur, not because priming and skill learning are based on inherently different mechanisms, but simply because the pursuit-rotor task involves a large motor component, whereas the word-completion task does not. Similar considerations apply to a variety of other neuropsychological data that have been interpreted as evidence of processing dissociations between skill learning and repetition priming. For instance, Nissen, Knopman, and Schacter (1987) found that administration of an amnesic agent to unimpaired participants resulted in impaired priming in a word-fragment completion task, but unimpaired skill learning in a serial reaction time task. Here, too, however, the dissociation could reflect differences in the processing requirements of the two tasks rather than the existence of different mechanisms for skill learning and repetition priming. For a similar but more detailed analysis of neuropsychological dissociations, the reader is referred to Poldrack et al. (1999).

Our framework also provides perspective on recent neuroimaging results that appear to suggest dissociations between skill learning and repetition priming. A number of neuroimaging studies have indicated that skill learning is characterized by increases in the intensity or spatial extent of brain activation (e.g., Grafton, Hazeltine, & Ivry, 1995; Grafton et al., 1992; Hazeltine, Grafton, & Ivry, 1997; Karni et al., 1995). In contrast, in a number of neuroimaging studies, repetition priming has been accompanied by decreases in activation (e.g., Buckner et al., 1995; Demb et al., 1995; Squire et al., 1992; Wagner, Desmond, Demb, Glover, & Gabrieli, 1997). Further information may be found in a number of recent reviews of the neuroimaging literature on repetition priming (Schacter & Buckner, 1998; Schacter, Wagner, & Buckner, 2000; Wiggs & Martin, 1998).

Our framework suggests that these findings need to be interpreted with caution. As Poldrack, Desmond, Glover, and Gabrieli (1998) have noted, in these neuroimaging studies, skill learning has been examined primarily in tasks involving motor processes, such as finger-tapping (e.g., Karni et al., 1995), and repetition priming has been examined primarily in tasks involving higher level cognitive processing, such as word-stem completion (e.g., Buckner et al., 1995). Our earlier cautions about making inferences about underlying relationships between skill learning and repetition priming that are based on performance in different tasks in neuropsychological studies therefore apply to the neuroimaging results as well.

In one of the few neuroimaging studies to date that has examined skill learning and repetition priming in the same task, participants performed lexical decision on mirror-reversed text stimuli (which were either novel or practiced), and it was found that skill learning and repetition priming were both associated with patterns of both increase and decrease in activation (Poldrack et al., 1998), confirming the idea that the dissociation between increasing activation for skill learning versus decreasing activation for repetition priming is less clear cut than it might appear to be in other neuroimaging studies. The question remains, nevertheless, of why the patterns of activation were not identical for skill learning and repetition priming in the Poldrack et al. (1998) study, given that both phenomena were studied in exactly the same task. There are at least two factors that need to be kept in mind here. The first is that the effect of repetition of nonwords in lexical decision tasks is not always easy to interpret: On the one hand, repetition of nonwords should lead to facilitation at some levels of processing (i.e., in some of the "processors" deployed in the task); on the other hand, at other levels of processing, there may be interference, as the repeating nonwords become more familiar and hence more susceptible to being accepted as real words. This point is borne out by Kirsner and Speelman (1996), who found that lexical decision accuracy for the repeating nonwords actually decreased as a function of experimental practice. As the response task used in the Poldrack et al. (1998) study was lexical decision, this makes the results somewhat difficult to interpret.

There is also a more general consideration, however, arising from our theoretical framework. In that framework, as we outlined above, any task is performed by a collection of processors, and facilitation of processing may occur in any subset of those processors. Any overall skill learning effect that is observed in the task is a manifestation of the sum of facilitation in various processors. Similarly, any overall repetition priming effect that is observed in the task is a manifestation of the sum of facilitation in various processors. However, the loci of facilitation are not necessarily the same for the skill learning and repetition priming effects (although, in our view, the underlying facilitation mechanism—incremental tuning—is the same across all processors).

To see this, consider a task in which a participant is presented visually with familiar words, one at a time, and must judge how many syllables each word has. One processing component that is required for this task is the processor or ensemble of processors that underlie the reading of single words. Another component is the processor or ensemble of processors that must be deployed to make the syllable judgment; this component is much more task specific than is the first one. The first processor (for reading single words) is highly overlearned (in a literate adult), and there is little potential for generalized facilitation in this processing component. Any skill learning observed in the overall task is therefore unlikely to arise from facilitation in this component of the overall system performing the task; its more likely locus is in the second, more task-specific processing ensemble, which is not overlearned. The locus of any repetition priming that is observed in the overall task may, however, be more evenly distributed across the two processing components. This is because, although there may be little possibility for generalized facilitation in the single-word-reading processing ensemble, there is likely to be greater possibility of facilitation toward specific items. (The analogy with connectionist networks may again be useful here: The network is at such a high level of performance that adjustment of connection weights in response to specific stimuli will facilitate performance on those specific stimuli on a subsequent exposure, but will not contribute greatly to generalized facilitation on novel stimuli. We will return to a more detailed discussion of these points later in the article in the sections entitled *One View of Procedural Memory and Learning* and *Multiple Levels of Skill Learning and Repetition Priming*).

According to our theoretical framework, then, although the underlying mechanism of facilitation—procedural learning—is the same in all the processing components of a task, any observed skill learning and repetition priming effects may have differing loci within the system of processors deployed in the task. What this means, however, is that in a neuroimaging study, skill learning and repetition priming may be associated with nonidentical patterns of activation, even when these effects are examined in a single task. We would therefore caution against the easy interpretation of differing patterns of activation as reflecting different underlying mechanisms.

In the *Necessary and Sufficient Conditions* subsection, we will discuss a number of predictions that follow from our theoretical framework as well as specific patterns of data that our theory would be unable to account for. However, to our knowledge, none of these patterns of data has ever been observed. In summary, then, the theoretical analysis we have offered of skill learning and repetition priming makes it clear that the various arguments that have been made for dissociations between the mechanisms underlying skill learning and repetition priming are either invalid (e.g., with respect to the implications of power functions or correlations) or inconclusive (with respect to patterns of increase and decrease). Dissociations that are suggested by the available neuropsychological and neuroimaging data are also problematic as evidence of separate processing mechanisms, as we have attempted to show. In fact, as far as we are aware, there is no conclusive evidence that skill learning and repetition priming arise from separate mechanisms. Of course, this does not in itself show that a single mechanism does underlie skill learning and repetition priming. In the second part of this article, we therefore turn to an examination of whether skill learning and repetition priming can indeed arise from a single mechanism.

Computational Analysis

We begin with an informal presentation of our view of procedural memory and describe the nature of the proposed underlying mechanism. We discuss how, in this view of procedural memory, repetition priming and skill learning are manifestations of the same mechanism. We then present a computational model of performance in the digit-entering task that incorporates this theory of procedural memory. We show that this model exhibits both skill learning and repetition priming, even though it consists of only a single learning mechanism. We show, moreover, that the performance of this model exhibits the same patterns of dissociation that have led to the presumption of different underlying mechanisms in skill learning and repetition priming.

One View of Procedural Memory and Learning

In our view, the performance of any task can be conceptualized as a process of transduction between representations. For example, the process of verbally expressing an idea or thought requires transduction from a conceptual representation of the thought-idea to the representations of an appropriate sequence of oral-motor commands. In this case, as in many others, there may be several intermediate levels of representation. To transduce one representation into another, there must be a transducer between those levels of representation. Given that there may be several intermediate levels of representation, there may in general be several transducers that underlie the performance of a given task. All the transducers that are involved in performance of a task undergo incremental adjustment, or tuning, each time the task is performed (i.e., each time a particular input is processed). Improvement in performance on a task arises from improvement in the effectiveness of the transducers that are involved in performance of that task. The improvement in effectiveness of transducers arises from the cumulative effect of the incremental tuning process, operating over many instances of performance of the task. This incremental tuning is, in our view, the basis of procedural memory and learning.

This view does not claim that there is a single procedural memory system. On the contrary, procedural memory and learning in a particular task are viewed as being situated in whatever processors are required for that task. Thus the locus of procedural learning-memory may be widely different in tasks that draw primarily on, say, motor versus perceptual processing. Similarly, for a task that requires a variety of processing systems (e.g., motor as well as perceptual), procedural learning may occur in several or all of these processing systems. Thus we do not claim that there is a single or specific set of processing elements constituting a procedural memory "system." What we claim, rather, is that the nature of the procedural learning mechanism is the same in all these cases. It is based on the incremental tuning of processors, whatever or wherever those processors (i.e., transducers) may be.

Improvement in performance in a task may be manifested in (a) improved performance on a specific stimulus or set of stimuli; (b) improved performance on stimuli in general (i.e., improved generalization); or (c) both of the above. Skill learning and repetition priming are therefore simply different manifestations of improvement in the effectiveness of the tranducers on which a particular task depends. Skill learning is manifested in improved generalization ability. Repetition priming is manifested in improved performance on specific items, following exposure to those specific items. Both of these effects are an inherent part of the way the system works. One way of stating this view is to say that skill learning and repetition priming arise from a single mechanism. A more accurate statement of this view is that skill learning and repetition priming are merely the epiphenomenal outcomes of repeated performance of any task. The "mechanism" that they arise from is not a special process or mechanism that provides for skill learning or repetition priming. It is merely the ongoing incremental learning inherent to the operation of the system itself, where "system" refers to the collection of transducers that are involved in performance of the particular task.

To further develop these ideas, we need to introduce a distinction between transductions and mappings. For any two levels of representation, we can define various possible sets of transductions. Each set of transductions identifies a set of particular representations at the first level and specifies which particular representation at the second level each of those first level representations maps onto. We will refer to such a set of transductions as a *mapping* between the two levels of representation. Note that a mapping is an abstractly defined set of transductions. A *transducer* is a physical device that approximates a particular mapping. The transducer comes to approximate the mapping by being exposed to sample input–output pairs that are drawn from the mapping (which defines the population of all possible pairs) and by undergoing modification as a result of each exposure. This tuning (i.e., procedural learning) occurs in such a way as to make the transducers more effective in approximating a particular mapping. That is, procedural learning enables the transducers to encode knowledge of the mapping. Procedural memory is just a demonstration of that knowledge, as revealed in performance of the task. The distinction between procedural memory and learning is therefore blurred.

This view of procedural learning and memory can be operationalized in terms of the framework offered by connectionist networks. The system of weighted connections between two "layers" of a connectionist network is an example of a transducer. The process whereby a representation at the first layer evokes a representation at the second layer in such a network is a process of transduction from the first to the second level of representation. In the performance of any task, several levels of transduction (and hence several transducers) will usually be required. We assume that representations at each level are distributed over many processing units, so that two similar stimuli are represented at a particular level by similar (i.e., overlapping) patterns.

In this connectionist formulation, tuning of the transducers in a given task corresponds to incremental weight adjustment in the connections between layers. The representation evoked at Layer 2 by a representation at Layer 1 can generally be expected to diverge to some degree from the appropriate, or target Layer 2 representation, this divergence being the error. The incremental adjustment of connection weights can enable the particular input representation at the first layer to evoke a closer approximation of the target representation at the second layer the next time that particular transduction needs to be performed. Thus improvement in effectiveness of the transducers arises from the cumulative effect of incremental weight changes over many instances of performance of the task. Improved performance on stimuli in general corresponds to improvement in the effectiveness of the overall transduction such that there is improved generalization ability. Such generalization is possible if the system employs distributed representations, so that similar stimuli have overlapping representations and therefore can share connection weights. Improvement in performance on specific stimuli can arise as a result of weight change that follows exposure to those specific stimuli. These ideas are closely related to those put forth by McClelland and Rumelhart (1985) regarding the representation of general and specific information in a distributed network with a single set of weighted connections; they are also implicit in our own previous work (Cohen & Eichenbaum, 1993).

Having presented our view of procedural memory, we can now consider how it applies to a specific task, the digit-entering task, which we described briefly in the *Implications of Correlations* subsection. In the version of the task we discussed, five-digit number strings (e.g., *49385*) were presented individually to participants, who entered these number strings using a numeric keypad (Poldrack et al., 1999, Experiment 1). In each block, some number strings appeared multiple times (repeating items) while other items appeared only once (unique items). We can conceptualize this task as requiring transduction from the visual representation of a five-digit number string to the representation of an appropriate sequence of finger-movement commands. The existence of skill learning in the digit-entering task indicates that the transducer(s) underlying this transduction improve(s) in effectiveness, leading to better generalization ability. This improvement arises from the effects of incremental tuning over many instances of task practice. The finding of repetition priming in the digit-entering task indicates that tuning in the transducer(s) also occurs for specific stimuli that are repeatedly encountered, and that this leads to better performance on these specific stimuli than on stimuli in general. This seems very reasonable: If the effectiveness of the overall transduction improves as a result of exposure to unique stimuli, it follows that the transductions for the repeating stimuli should be tuned to an even greater extent because of the repeated exposure to these specific stimuli.

If these ideas are correct, then it should be possible to construct a computational model that can simulate behavioral performance in the digit-entering task. The remainder of this section develops such a model. The model incorporates our view of procedural learning and memory and adopts the computational framework of PDP (McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986). It should be noted that we have taken a similar approach in previous work (Poldrack et al., 1999), constructing computational models to demonstrate that skill learning and repetition priming can arise from a single mechanism. However, the results reported here differ from and extend the previous work in important ways. First, the computational model that is developed here is an explicit psychological model of performance in a specific task in which both skill learning and repetition priming have been studied; in contrast, our earlier work presented abstract computational models that did not represent performance of any specific task. Second, we use the present model to simulate a complex set of empirical results from an experimental task; this is to our knowledge the first computational account of specific experimental data on skill learning and repetition priming within a single domain. Third, we will show that empirically observed patterns of dissociation between repetition priming and skill learning arise naturally in the present model, without manipulation of any of its parameters. For these reasons, we believe that the present work constitutes a more compelling demonstration of the relatedness of skill learning and repetition priming than has previously been made and therefore represents a significant advance in theoretical development.

The Digit-Entering Task

Prior to describing our computational model, we need to provide a more complete description of the digit-entering task, which we have so far described only in general terms. We will focus on the version of the digit-entering task reported by Poldrack et al. (1999). In this task, five-digit number strings (e.g., *49385*) were presented individually on a computer display to participants, who entered these number strings using a numeric keypad. No visual feedback was provided for the response. After entering the number string, the participant pressed the “Enter” key, which triggered the next trial after a 1-s intertrial interval. Participants were instructed not to correct errors; they were told to press the “Enter” key if they made an error, to proceed to the next trial. In each session, some number strings appeared multiple times (repeating items), while other items appeared only once (unique items).

Training stimuli were chosen from the set of all possible five-digit number strings by placing three constraints on them. The first constraint was that only the digits 1–9 could appear in the number strings (i.e., the digit 0 could not appear). This reduced the set of permissible five-digit number strings to 59,049 possible strings. Second, digits could not repeat immediately (e.g., “44” never appeared as part of a number string). Third, five-digit number strings were constrained to obey certain first-order transitions between digits, as follows: For each of the digits 1–9, there are eight possible (different) digits that can follow it. For each digit, four of these eight possible transitions were chosen at random. The set of 36 chosen transitions (4 transitions for each of the nine digits) composed one full transition rule set. There was also a complementary transition rule set composed of the other 36 transitions (4 complementary transitions, for each of the nine digits. (See Figure 8 for a diagram of such transition rule sets for the digit 1). Five-digit number strings were constrained to obey one or other of these transition-rule sets. One half of the participants in each experiment were trained on five-digit number strings that followed one transition rule set, and the other half were trained on items following the complementary rule set.

The experiment (Poldrack et al., 1999, Experiment 1) consisted of three sessions, each consisting of 12 blocks. The structure of the experiment is summarized in Table 2. Forty-eight five-digit strings were presented in each block. However, the composition of the 48 stimuli in a block varied across sessions. In Session 1, each block consisted of 12 repeating stimuli that appeared once in every block throughout Session 1 and also throughout the experiment (S1 Repeats). The remaining 36 stimuli in each block of Session 1 were unique stimuli, that is, stimuli that appeared once, in only that one block, and that did not appear in any other block in the experiment (S1 Uniques). In Session 2, each block consisted of the 12 repeating stimuli from Session 1, 12 new repeating stimuli that appeared in each block in Session 2 and through the rest of the experiment (S2 Repeats), and 24 unique stimuli (S2 Uniques). In Session 3, each block consisted of the 12 Session 1 repeating stimuli, the 12 Session 2 repeating stimuli, and 12 unique stimuli (S3 Uniques). In addition, each block in Session 3 contained a further 12 unique stimuli that did not conform to the transition rule structure of all the other stimuli (New Rule Uniques). That is, all the stimuli in all three sessions of the experiment conformed to a particular set of transition rules except for the Session 3 New Rule Uniques, which conformed to the complementary rule set.

Response times were separated into two components: latency and interkeystroke interval (IKI). Latency was the time from presentation of the stimulus to the first keypress. IKI was the average interval between subsequent digit keystrokes (i.e., keystrokes for digits 2–5, but not the keystroke for the “Enter” key). Poldrack et al. (1999) reported both latency and IKI and found the skill learning and repetition priming results to be similar using either measure. In this article, we summarize only the IKI data of Poldrack et al. (1999, Experiment 1), as these are the results we

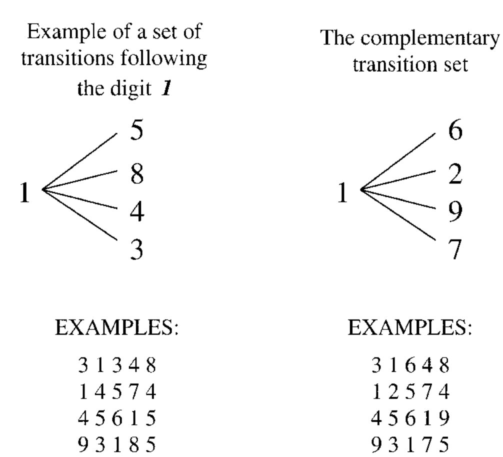


Figure 8. Example of transition rules used in the digit-entering task of Poldrack et al. (1999).

Table 2
Structure of the Digit-Entering Task in Poldrack et al. (1999, Experiment 1)

Session	Trial type	Number of stimuli per block
1	S1 repeats	12
	S1 uniques	36
2	S1 repeats	12
	S2 repeats	12
	S2 uniques	24
3	S1 repeats	12
	S2 repeats	12
	S3 uniques	12
	New rule uniques	12

Note. S = session.

will target in our simulations. The reason for this modeling choice will be discussed following description of the model. The IKI data we present here include participants' correct as well as incorrect responses.

This experiment was designed to address several questions. First, how does the amount of priming vary with the level of overall skill? This, of course, is the basic question that has been asked in investigations of the relationship between skill learning and repetition priming. In Experiment 1, data pertinent to this question were provided by comparing performance on S1 Repeats throughout the three sessions with performance on the Old Rule Unique patterns through the three sessions (which consist of the S1 Uniques, S2 Uniques, and S3 Uniques, but exclude the New Rule Uniques that were presented in Session 3).

A second question asked in the experiment was whether the priming observed for repeating items can be obtained at different levels of skill learning. This question was addressed by the introduction of new repeating items in Session 2 (the S2 Repeats). At the point of their introduction in the first block of Session 2, the S2 Repeat patterns were unique stimuli for the participant. Performance on these stimuli should therefore be no different from performance on S2 Uniques in the first block of Session 2. Any divergence in performance on subsequent blocks (between the S2 Repeat stimuli and the S2 Unique stimuli) would therefore indicate that repetition priming can occur even when a considerable amount of skill learning has already occurred (it was expected that significant skill learning would have occurred by the beginning of Session 2 as a result of practice in Session 1).

A third question, of lesser importance for present purposes, was whether the lower level statistical regularities of the stimuli are represented as part of skill learning. This question was addressed by comparing performance on new items from the studied transition rule set with performance on new items from the complementary transition rule set, that is, by comparing performance on S3 Uniques and S3 New Rule Uniques. To the extent that there was negative transfer to the complementary rule set items, this would indicate that statistical regularities were learned in the task.

Figure 9 illustrates pertinent results from Poldrack et al. (1999, Experiment 1). The horizontal axis shows blocks of practice, and the vertical axis shows participants' performance as measured by mean IKI per block. The four curves in the upper part of the figure show (from the top downward) performance on S3 New Rule Unique stimuli, performance on Old Rule Unique stimuli, performance on S2 Repeat stimuli, and performance on S1 Repeat stimuli. Skill learning was measured as the improvement in performance on Old Rule Uniques over the course of the experiment. That is, skill learning at block *n* is the difference between performance on Old Rule Uniques at Block 1 and performance on Old Rule Uniques at Block *n*. This measure is plotted as the upper of the two curves in the lower part of the figure. Repetition priming at Block *n* was measured as the difference between performance on Old Rule Unique stimuli and performance on S1 Repeat stimuli, both measured at Block *n*. This measure is plotted as the lower of the two curves in the lower part of the figure.

The following effects are noteworthy. First, there is marked improvement in performance on both repeating stimuli and Old Rule Unique stimuli across blocks in Figure 9. Second, there is a clear increase in skill learning across blocks, seen in the decrease in IKI for Old Rule Unique items across blocks and in the plot of skill learning in the lower part of the figure. There clearly also is repetition priming, as shown by the difference between performance on Old Rule Unique stimuli and performance on S1 Repeat stimuli at each block and by the plot of repetition priming in the lower part of the figure. Thus, the results shown in Figure 9 clearly indicate the presence of both phenomena of interest, although repetition priming does not appear to increase as much as skill learning does across blocks. Third, repetition priming can be observed at different levels of skill learning. This can be seen in

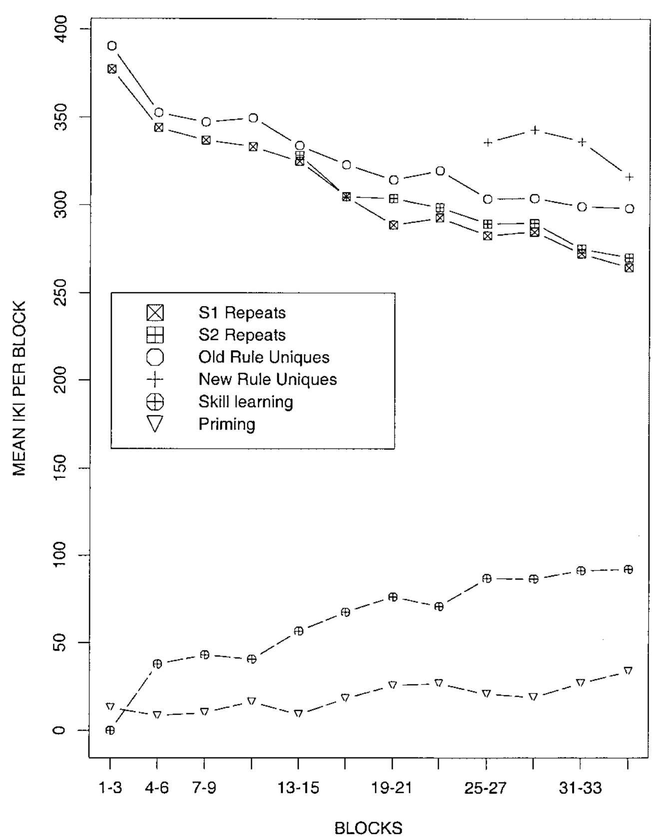


Figure 9. Key results from the digit-entering task reported in Poldrack et al. (1999, Experiment 1). IKI = interkeystroke interval; S = session.

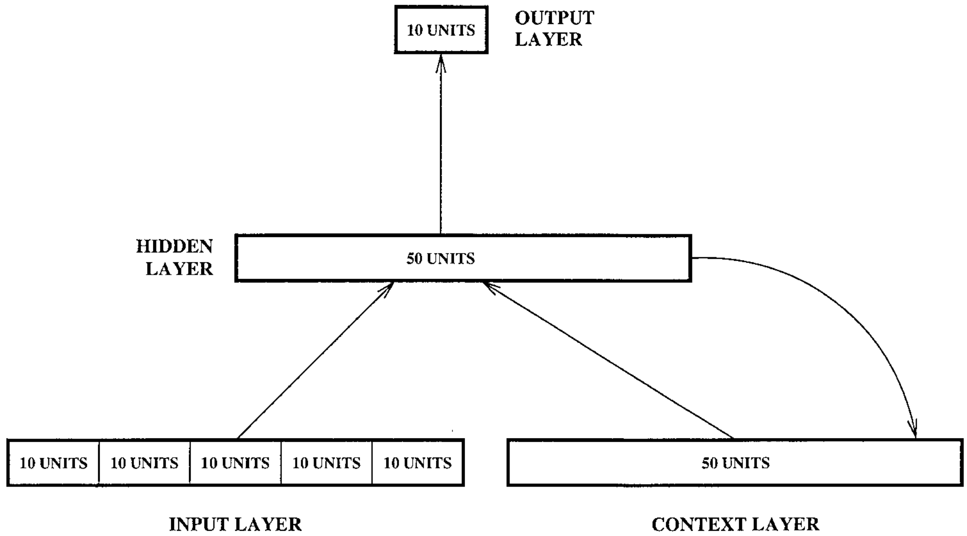


Figure 10. Architecture of the model of the digit-entering task.

the divergence in performance between S2 Repeat stimuli and Old Rule Unique stimuli, following introduction of the S2 Repeat stimuli in Block 13. Fourth, there is negative transfer to the New Rule Unique stimuli introduced in Session 3. The only difference between the Old Rule Unique stimuli and New Rule Unique Stimuli is in the familiarity of the specific two-digit transitions they incorporate. The negative transfer effect therefore indicates that participants learned something about the two-digit transition structure of the Old Rule Unique stimuli. That is, part of what they learned were subitem regularities within the stimuli.

A Computational Model of the Digit-Entering Task

Description of the model. In each trial in the digit-entering task of Poldrack et al. (1999), participants were presented with a five-digit number string, which they were required to convert into a sequence of keyboard strokes that represented the sequence of digits composing the five-digit number string. Our model's task was formulated to have the same structure: The input to the model was a representation of a five-digit number string, and the model's task was to produce a sequence of outputs that represented the sequence of digits in the input. Figure 10 illustrates the architecture of the model, which is an adaptation of simple recurrent network architectures that have been studied by several researchers (e.g., Cleeremans, Servan-Schreiber, & McClelland, 1989; Elman, 1990; Jordan, 1986).

The input was represented as a pattern of activation over five banks of units that constituted the Input Layer. Each of these banks was composed of 10 units, which represented the digits 0–9, with each unit representing a specific digit. Thus the five-digit string "49683" was represented as a pattern of activation in which the 5th unit in the first bank of units was active, representing the occurrence of "4" as the first digit in the string; the 10th unit in the second bank of units was active, representing the occurrence of "9" as the second digit in the string; the 7th unit in the third bank was active, representing the occurrence of "6" in third position; the 9th unit in the fourth bank was active, representing the occurrence of "9" as the fourth digit; and the 4th unit in the fifth bank was active, representing the occurrence of "3" as the fifth digit. All other units in the 50-unit Input Layer were inactive.

The Output Layer consisted of a single bank of 10 units, representing the digits 0–9, with each output unit representing a specific digit. When presented with an input pattern of the kind described above, the network's task was to "spell out" the input as a sequence of activations at the Output Layer. For the input "49683", the network had to activate, in sequence, the Output Layer units representing the digits "4", "9", "6", "8", and "3".

At the first step in processing this input, activations from the Input Layer and from the Context Layer were propagated forward to the Hidden Layer, and activations from the Hidden Layer were in turn propagated forward to the Output Layer. This resulted in some pattern of activation at the Output Layer. An error signal was generated, representing the distance of this actual activation pattern from the target pattern, in which the output unit representing "4" should be maximally activated and all other output units inactive.[5] The error signal was then used to adjust the strength of connection weights throughout the system to reduce error via the "back propagation" of error (Rumelhart et al., 1986).

[5] The rationale for assuming the availability of a target pattern is as follows. A naive human participant performing the digit-entering task can be assumed to know, at a cognitive level, what the correct sequence of key presses should be for any given digit string. We assume that this cognitive representation can serve as a teaching target for the subsystem that actually executes a sequence of key presses. Our model and simulations represent the execution system and accordingly, we assume the availability of a target pattern. These assumptions are analogous to those made in similar models. For instance, in Seidenberg and McClelland's (1989) model of word naming, it was assumed that a target pattern can be self-generated on the basis of the learner's prior knowledge of the orthography and phonology of words. See Jordan and Rumelhart (1992) for a more detailed discussion of "teaching signals" and their availability to connectionist models.

At the next step in processing the input, the activation of Hidden Layer units was copied to the Context Layer (Elman, 1990), and the network was now expected to produce as output the second digit in the input string. Note that the input pattern that represented the digit string *49683* was still present. As in the first step of processing, the actual output produced by the network was compared with the target output for the second digit and connection weight strengths adjusted to reduce the magnitude of error. Processing continued in this way for three further time steps, at which the network's task was to output the third, fourth, and fifth digits, respectively, of the input. At the end of the fifth time step, the Context Layer was cleared, and the next five-digit input string was presented for the network to "spell out."

It is worth clarifying some representational decisions that were incorporated in the model, particularly as regards choices between distributed and localist representations. The crucial characteristic of distributed representations is that they provide a basis for the representation of similarity between different stimuli. The use of a distributed scheme for a particular level of representation is appropriate, therefore, when it is theoretically important that there be a basis for generalization from one stimulus to another similar stimulus at that level of representation.

In our account of the digit-entering task, the system must be capable of processing the novel five-digit string *59378*, even if it has never previously encountered it, where "processing" means being able to produce the correct sequence of outputs when presented with a representation of the entire five-digit string at the input. A system will be able to process a novel string such as *59378* correctly if it has previously processed similar five-digit strings such as *59371*, *59376*, *42568*, and *31728* and can generalize from these similar previous strings. For the system to be able to generalize in such fashion, there must be a basis for representing the similarities between similar five-digit strings. This calls for the use of distributed representations at the input layer in our account of the digit-entering task. And in fact, our representations of five-digit stimuli at the input layer are distributed representations. For example, they provide a means of representing the similarity between the digit strings *59376* and *59378*: The patterns of activation that represent these strings are identical across the first four banks of input units and differ only in activation within the last bank of units.

We could have used a distributed scheme at the output layer also; for instance, one in which each digit was represented as a pattern of activation across a set of units that represented various visual features of digits. However, it was not critical for our theoretical account of the digit-entering task that the system be able to generalize from one digit to another; the use of a distributed representation would therefore not have affected the results. For this reason, we did not incorporate distributed representations at the output level. The same is true of our decision not to use distributed representations for individual digits at the input level.

Description of simulations. The model was used to simulate performance in the digit-entering task of Poldrack et al. (1999, Experiment 1). Sixteen participants took part in that experiment. One complete simulated "replication," accordingly, comprised 16 runs of the model. Each run of the model consisted of three parts. First, weights in the system were initialized to random values, corresponding to variation between participants. Second, the model was trained on 1,000 five-digit strings, which were randomly chosen from the set of all possible five-digit strings, but which excluded strings that would appear in the actual experimental simulations. This training consisted of four epochs, that is, four cycles of presentation of the 1,000-stimulus set. The model's task at each five-digit stimulus presentation was to spell out the sequence of digits composing the string. This second step of pretraining was intended to provide the model with a preexperimental level of skill approximately equivalent to that of participants entering the experiment. The rationale was that human participants enter the experiment with significant practice in translating digit strings into output sequences of digits, whether on a keyboard or in writing. At the end of preexperimental training, the model was correctly able to spell out 85.3% of novel five-digit inputs, which corresponded quite well with the 86.98% accuracy of human participants on unique stimuli in the first block of Poldrack et al.'s (1999) digit-entering task.

The third step consisted of the actual experimental phase of the simulation run. The structure of this third phase precisely mirrored the structure of the digit-entering task in Poldrack et al. (1999, Experiment 1), which was summarized in Table 2. Thus, in the first epoch of a simulation, the model was presented with 48 stimuli (five-digit strings) that obeyed a particular transition rule structure. Twelve of these stimuli were to repeat throughout the simulation (S1 Repeats), whereas the other 36 stimuli would not be presented again during the simulation (S1 Uniques). As explained in the previous section, the model's task at each five-digit stimulus presentation was to spell out the sequence of digits composing the string, just as in the experiment with human participants. Thus Epoch 1 in the simulation corresponded to Block 1 in the experiment. In the second epoch of a simulation, the 12 S1 Repeat stimuli were presented again, and 36 novel stimuli were also presented. Thus Epoch 2 corresponded to Block 2. In analogous fashion, each epoch in the simulation mirrored the corresponding block of the experiment. It may be worth noting that the model's task in response to presentation of a five-digit string was the same during Step 2 (pretraining) and during Step 3 (experimental simulation). The only difference was that the presentation of stimuli during Step 3 had a carefully controlled structure, just as in the digit-entering task of Poldrack et al. (1999, Experiment 1).

In Experiment 1 of Poldrack et al. (1999), the specific stimuli in each category (S1 Repeats, S1 Uniques, etc.) were carefully counterbalanced across the 16 participants. In simulations, the model was run once with each of these actual sets of stimuli. Thus one simulation run consisted of (a) random weight initialization; (b) pretraining for 4 epochs; and (c) simulation of the 36 Blocks of Experiment 1, using one of the 16 stimulus sets actually presented to participants. A set of 16 such runs constituted one simulated replication of Experiment 1. Twenty such replications were run and the results averaged.[6]

Simulation results. The latency and IKI measures that were derived from the digit-entering task represent somewhat different components of the response to a particular five-digit stimulus. In

[6] All simulations were run using the **bp** program of McClelland and Rumelhart (1988). Weights were initialized with values distributed uniformly between −0.5 and +0.5 and were updated after every pattern presentation. The momentum parameter was set to 0.9. The learning rate parameter was set to 0.0005 for simulations of the experiment (Stage 3 of the simulations), reflecting our view that procedural learning consists of

particular, the latency measure is likely to include time that is required for planning of the five-digit sequence as a whole, while the IKI measure is likely more indicative of the execution of the individual components of the five-digit sequence (as previously noted by Fendrich et al., 1991, p. 139). In the computational model described above, there is no "planning" process during which the plan is assembled; rather, the entire plan is presented to the model in fully assembled form. What the model does is to execute this plan, one digit at a time. Thus, our computational model has no plausible analogue of the planning latency component of processing. However, it does have an analogue of the execution of individual keystrokes. For this reason, we view the model's behavior as affording a simulation of IKI data but not of latency data. (It is important to note, however, that this is a limitation of the particular implementation that we have chosen for the model and not an in-principle limitation of our approach. A model could be constructed in which assembly of the plan itself required processing in the model. In the interests of simplicity, we chose not to pursue such an implementation.)

To relate the patterns of activation that were produced by the model to the experimental IKI data, we used mean squared error (*MSE*) per digit, which was the mean squared difference between the target output pattern and the actual output pattern for every digit.[7] For example, the model's performance on Unique Stimuli in Block 1 (Epoch 1) of the simulations was measured by the squared error per digit, averaged over all digits presented in Block 1, averaged over the correct as well as incorrect responses of all 16 simulated participants, averaged over all 20 replications. We found that this measure was linearly related to the IKI measure that was used in the behavioral experiments. The simulated IKI measure we report here is derived from *MSE* by the equation $IKI_{\text{simulated}} = (MSE * 10{,}000)/3$. That is, we needed only a linear model with a single parameter to approximate the magnitude of the behavioral IKI measure.

Figure 11a redisplays the results from Poldrack et al. (1999, Experiment 1). Figure 11b shows the corresponding simulation results. As can be seen, the simulations provide a good fit to the data and exhibit all of the phenomena that are characteristic of the empirical results: (a) improvement in performance on both unique and repeating stimuli across blocks; (b) the existence of both skill learning and repetition priming; (c) the effects of repetition as observed even after significant skill learning has occurred, as shown by performance on the new repeating items (S2 Repeats) introduced in Session 2; and (d) the negative transfer to stimuli that follow the complementary transition rule structure in Session 3.

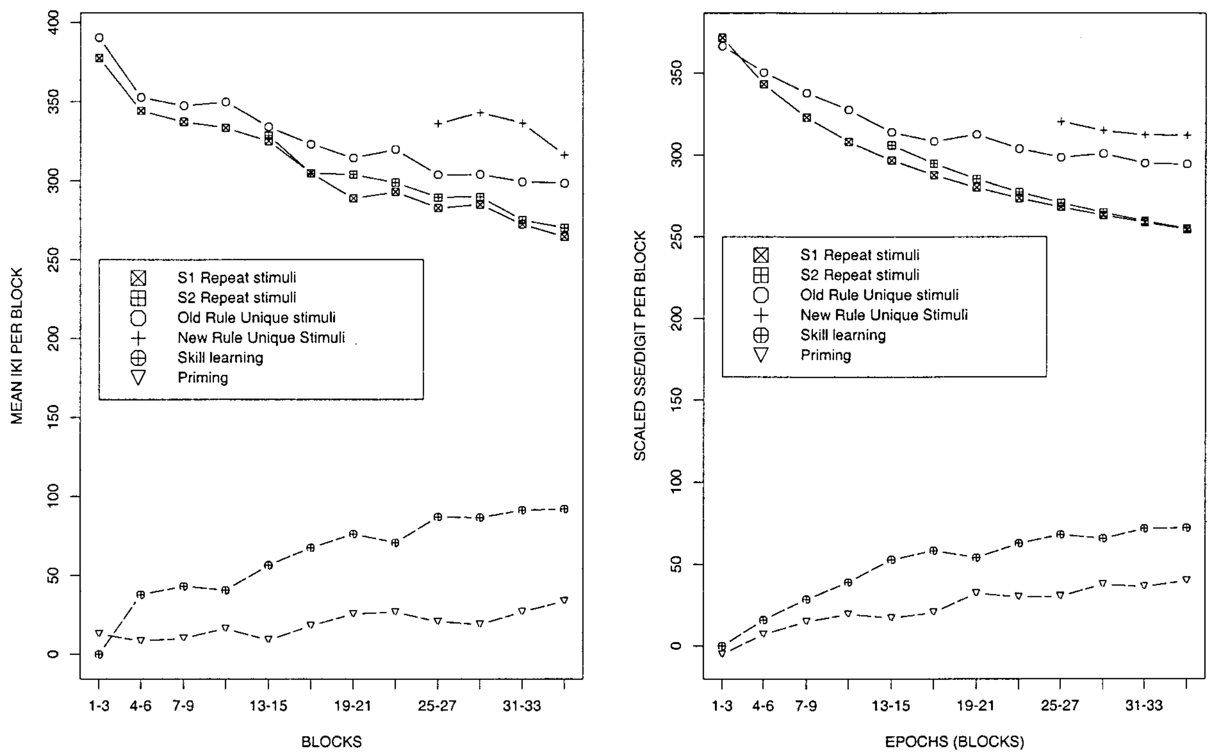
very gradual and incremental tuning or weight change. During pretraining (Stage 2 of the simulations), we used a higher learning rate of 0.085. This is because we make no claims about the correspondence of our pretraining regimen (four presentations of 1,000 pretraining stimuli) with the actual extent or nature of human participants' previous exposure to "spelling out" digit strings. Rather, pretraining simulations aimed merely to bring the model to approximately the same preexperimental level of "spelling out" performance as that exhibited by human participants at the start of the digit-entering task. We therefore adopted the computational expedient of using a higher learning rate to speed up the pretraining simulations.

These results demonstrate that the skill learning and repetition priming that are observed in the digit-entering task can be accounted for in a model that consists of a single mechanism. The fact that this single-mechanism model provides a good fit to a complex pattern of empirical data suggests that this demonstration deserves to be taken seriously.

It is further noteworthy that this one-mechanism model exhibits some of the very dissociations between skill learning and repetition priming that have been cited in the literature as evidence of dual underlying mechanisms. Let us first examine correlations between skill learning and repetition priming. We have already established that a lack of correlation between skill learning and repetition priming has no theoretical implication with regard to the relatedness or unrelatedness of underlying mechanisms (see the *Implications of Correlations* subsection). We also saw that the empirical results from a digit-entering task (Poldrack et al., 1999, Experiment 1) exhibit an inconsistent pattern of correlation between skill learning and repetition priming across blocks (Figure 6). What is the pattern of correlations in the model? Figure 12 shows the correlation between the skill learning and repetition priming measures at each epoch or block of the simulations just described. As can be seen, there is no consistent pattern of correlation, just as in the empirical data that is shown in Figure 6. The mean correlation between skill learning and repetition priming in the digit-entering task (Poldrack et al., 1999, Experiment 1) is $r = -.263$, averaged across all blocks. In the simulations, the mean correlation is $r = -.101$, averaged across all epochs. Even though they arise from a single underlying mechanism in the model, the two measures are uncorrelated. These results further strengthen our argument that empirical observation of an inconsistent pattern of correlations between the skill learning and repetition priming measures does not constitute evidence for different underlying mechanisms.

We also used our simulation results to examine another kind of dissociation that has been reported in the literature, namely, that between the form of functions for skill learning and repetition priming (e.g., Kirsner & Speelman, 1996). We have already established that the form of the repetition priming measure has no theoretical significance (see discussion in *The Form of Functions* subsection). However, it is of interest to examine whether such dissociations arise in the model. If they do, this would constitute a further and quite compelling dem-

[7] When a five-digit input is presented to the model, the spelling out of digits occurs in five time steps, with the production of each individual digit occurring in a single time step. That is, the production of a particular digit occurs in a single feedforward pass of activation from the Input and Context Layers, through the Hidden Layer, to the Output Layer. The production of each digit is therefore not a temporally extended process. It has been shown, however, that activations that are computed in such a manner correspond to the asymptotic activations that would be reached in a cascaded system (Cohen et al., 1990; McClelland & Rumelhart, 1988). It has also been shown that patterns that asymptote at a relatively low level of error will reach a criterion level of accuracy relatively quickly (Cohen et al., 1990). The error measure that is associated with a particular output pattern can therefore be viewed as an analogue of reaction time—in this case, IKI—in a feedforward system (McClelland & Rumelhart, 1988; Seidenberg & McClelland, 1989).



(a) (b)

Figure 11. Comparison of experimental and simulation results. (a) Results from the digit-entering task of Poldrack et al. (1999, Experiment 1). (b) Simulation results. IKI = interkeystroke interval; S = session.

onstration that such dissociations do not constitute evidence of separate underlying mechanisms.

To examine this question, we fitted functions to three measures, using data from the experimental digit-entering task (Poldrack et al., 1999, Experiment 1) and from our simulations. The three measures were performance on unique stimuli, performance on repeating stimuli, and inverted repetition priming (IRP). IRP at Block *n* was computed as the difference between final block repetition priming and repetition priming at Block *n*. We used this measure (instead of repetition priming) for reasons of practical convenience. Power functions of the form $y = a + bP^{-c}$ (where P is practice) were fitted using PASTIS (Cousineau & Larochelle, 1997), which in turn uses STEPIT (Chandler, 1969).

Figure 13a displays data points and best-fit power functions for performance on unique stimuli and performance on repeating stimuli in the digit-entering task (Poldrack et al., 1999, Experiment 1). Parameters of the best-fit power functions are shown in the upper half of Table 3, along with two goodness-of-fit measures, root-mean-squared deviation (RMSD) and r^2, which show that these power functions provide a good fit to performance on unique stimuli and performance on repeating stimuli in the digit-entering task. For IRP, we found that a linear power function provided as good a fit in terms of RMSD and r^2 as did the best fitting nonlinear power function.[8] This best linear fit is also graphed in Figure 13a, and its parameters are shown in the upper half of Table 3. These fits to empirical data from the digit-entering task exhibit the same pattern as that obtained by Kirsner and Speelman (1996): Performance on unique stimuli and performance on repeating stimuli follow the nonlinear power law, but repetition priming is linear. Does the same pattern of results hold in our single-mechanism model?

Figure 13b displays data points and fitted functions for data from the digit-entering task simulations that are described in this section. For the simulations, as for the empirical data, we found that nonlinear power functions provided an excellent fit to the performance on unique stimuli and performance on repeating stimuli measures. For IRP, however (again, as with the empirical data), a linear power function provided as good a fit as did the best fitting nonlinear power function. The parameters of these fitted functions are shown in the lower half of Table 3, along with the goodness-of-fit measures. Even though these measures arise from a single learning mechanism in the model, they exhibit the same pattern of dissociation as in the digit-entering task empirical results, viz., that performance on unique stimuli and performance on repeating stimuli follow the nonlinear power law, but repetition

[8] Note that a linear function of the form $y = a - bP$ is just a special case of the generally nonlinear power function $y = a + bP^{-c}$, where $P =$ practice.

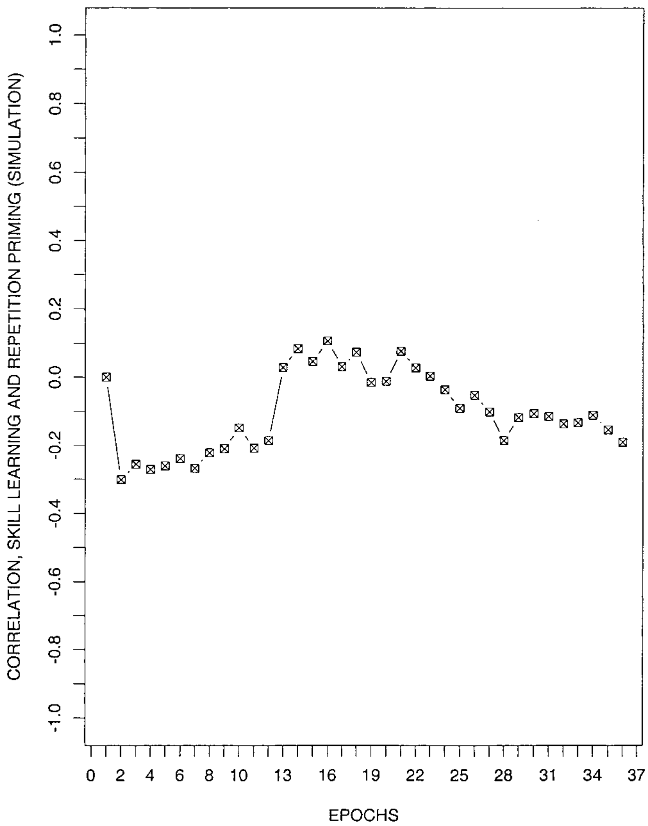


Figure 12. Correlation between skill learning and repetition priming (simulation results).

priming does not. This dissociation is similar, although not identical, to the pattern of dissociation found in Kirsner and Speelman's (1996) study. Clearly, such patterns of dissociation can arise from a single mechanism and therefore do not constitute evidence for separate underlying mechanisms.[9]

As a final point, note that the parameters of the functions that are fitted to the simulation results are not identical to the parameters of the functions that are fitted to the empirical results. This simply reflects the fact that the simulation results are not identical to the empirical results. Therefore, the best fitting functions are closely similar, but not identical, in the two cases. This raises the question of just how well the simulations model the empirical results. To make a quantitative evaluation of this, we compared the performance on unique stimuli curves from the empirical data, Figure 13a, and the simulation results, Figure 13b. The RMSD between the empirical and simulated performance on unique stimuli curves was 13.282 and $r^2 = .925$. The simulated performance on unique stimuli curve thus provides almost as good a fit to the empirical performance on unique stimuli curve as does the best fitting power function, whose RMSD and r^2 values are 5.974 and .950, respectively (Table 3). Similarly, the RMSD and r^2 values for the comparison between empirical and simulated performance on repeating stimuli curves, Figure 13a vs. 13b, are 16.771 and .953, respectively, which compare very favorably with the RMSD and r^2 values of 8.280 and .936, respectively, for the power function that provide the best fit to the empirical performance on repeating stimuli data (Table 3).

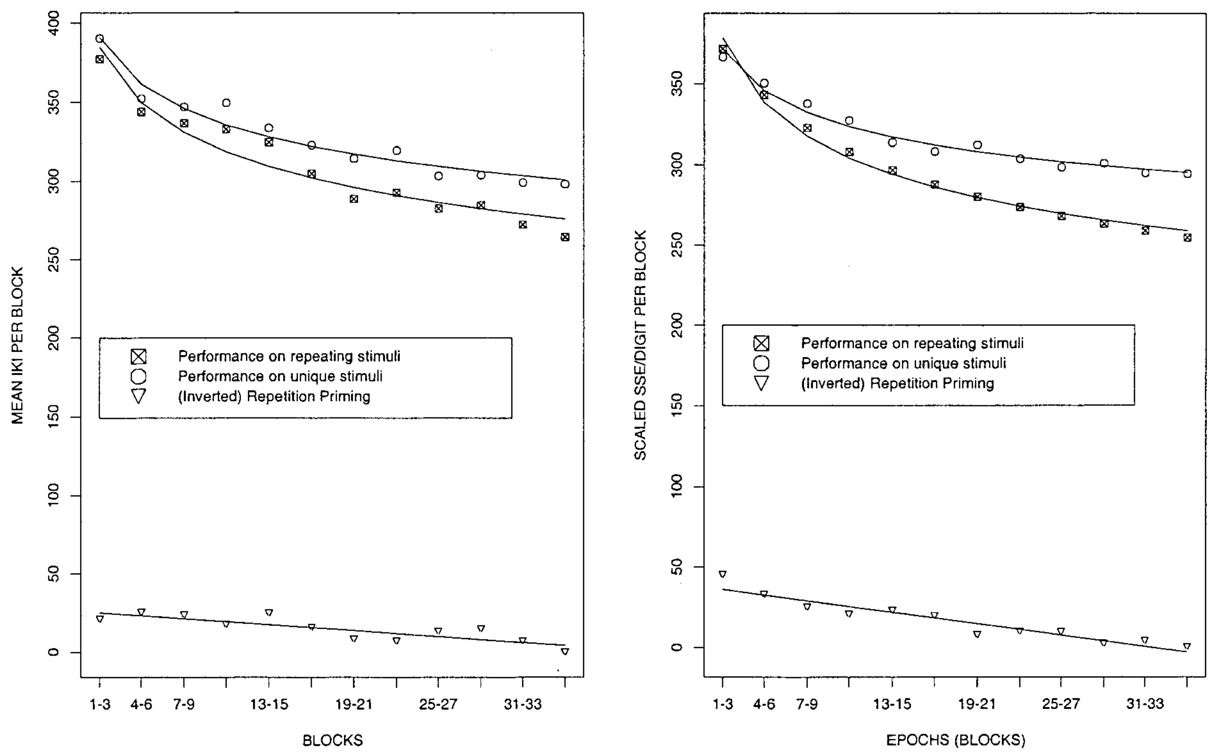
We take these various results as a strong demonstration that skill learning and repetition priming can indeed arise from a single underlying mechanism. Furthermore, as we have seen, even when arising from a single learning mechanism, they can exhibit the kinds of dissociation that have been taken as evidence of dual mechanisms. We believe that these demonstrations, together with the theoretical analyses that were presented in the first part of the article, provide strong evidence that the phenomena of skill learning and repetition priming are in fact outcomes of a single kind of learning mechanism.

Simulated Patterns of Increase and Decrease

We now return to our discussion in the *Patterns of Increase and Decrease* subsection. There we argued that all possible relationships between the skill learning and repetition priming measures were consistent with the view that the phenomena of skill learning and repetition priming arise from a single mechanism. Indication of the relatedness came from the close relations between the shape of the performance on unique stimuli and performance on repeating stimuli functions. There were, however, two cases where this was not immediately obvious from the shape of performance on unique stimuli and performance on repeating stimuli. One of these was Case II, in which repetition priming increases with practice, but there is no skill learning (see Figure 7). The other was Case IV, in which performance on unique stimuli improves with practice, but there is little or no improvement in performance on repeating stimuli, with the result that skill learning increases while repetition priming decreases. In this section, we demonstrate that each of these cases can arise from a single learning mechanism.

Repetition priming without skill learning. To demonstrate that the pattern of results in Case II can indeed arise from a single mechanism, we used our model for a further simulation. First, we pretrained the model for 15 epochs (instead of only 4 epochs, as in the simulations reported in the preceding section). The intent was to greatly improve the model's preexperimental level of skill at the basic task of spelling out digit strings. Following this preexperimental practice, we reran the simulations of Experiment 1, exactly as described in the previous section, and computed skill learning and repetition priming scores. The results are shown in Figure 14a. The results represent a pattern in which repetition priming increases in the absence of skill learning, and correspond to Case II in Figure 7. As a further check on what the simulation results represent, they can be compared with the results of Schwartz and Hashtroudi (1991, Experiment 1, word-fragment completion),

[9] In Kirsner and Speelman's (1996) results, skill learning increased as a function of practice and was best approximated by a curvilinear power function; in contrast, repetition priming did not increase as a function of practice and was best approximated by a (horizontal) linear function. The dissociation in the digit-entering task and in our simulation results is slightly different: Skill learning increases and is best approximated by a curvilinear power function, while repetition priming increases and is best approximated by an increasing linear function. The difference between the two patterns of dissociation is that in the Kirsner and Speelman study, skill learning increases as a function of practice, while repetition priming does not. However, as we explained in the *Patterns of Increase and Decrease* subsection, this pattern is not necessarily indicative of a processing dissociation either.



(a) (b)

Figure 13. Best-fit functions for performance on unique stimuli, performance on repeating stimuli, and inverted repetition priming. Points represent data, lines represent best-fit functions. (a) Fits to data from Poldrack et al. (1999, Experiment 1). (b) Fits to simulation results. IKI = interkeystroke interval.

which are redrawn in Figure 14b, and which represent a specific example of priming in the absence of skill learning. The skill learning and repetition priming functions from the simulation results are remarkably similar to the skill learning and repetition priming functions in Schwartz and Hashtroudi's (1991) results.[10] These simulations suggest that such a pattern of findings can indeed arise from a single mechanism.

Increasing skill learning with decreasing repetition priming. We next consider Case IV from Figure 7. In that scenario, performance on unique stimuli improves with practice, but there is little or no improvement in performance on repeating stimuli, so that skill learning increases while repetition priming decreases. Given the difference between the form of the performance on unique stimuli and performance on repeating stimuli functions, it is not entirely clear whether this pattern is actually consistent with a single underlying mechanism. We reasoned that such a pattern of results could arise if the repeating stimuli had been highly overlearned prior to experimental testing, that is, through preexperimental practice. During experimental testing in the digit-entering task, facilitation in performance on such repeating stimuli might then be small or nonexistent. It might still be possible, however, to obtain facilitation in generalized task performance (i.e., improvement in performance on unique stimuli) during experimental testing.

To test these intuitions, we performed a further simulation, using exactly the same architecture that we have described so far. We modified the set of 1,000 five-digit stimuli that were used for pretraining in the previous two simulations, by adding 10 stimuli that would later be the repeating stimuli in the experimental simulation. Twenty repetitions of each of these 10 stimuli were randomly interspersed throughout the pretraining corpus, which thus consisted of 1,200 stimuli (1,000 random stimuli, plus 20 repetitions of each of 10 to-be-repeating stimuli). The model was

[10] For completeness, we have included not only the skill learning and repetition priming functions, but also the performance on unique stimuli and performance on repeating stimuli functions, in Figure 14a and 14b. It should be kept in mind that improvement in performance as measured by reaction time is manifested as a numerical decrease with practice, as we see in the simulation results, and that improvement in performance as measured by proportion correct is manifested as a numerical increase with practice, as seen in Schwartz and Hashtroudi's (1991) results. What may appear to be different patterns for performance on unique stimuli and performance on repeating stimuli in the simulations compared with the results from Schwartz and Hashtroudi (Experiment 1, word-fragment completion) are therefore in fact quite similar patterns. Note the clear improvement in performance on repeating stimuli and the relative lack of improvement in performance on unique stimuli in both cases.

Table 3
Parameters of Best-Fit Functions for Digit-Entering Task Empirical and Simulation Results

Measure	Parameter *a*	Parameter *b*	Parameter *c*	Goodness of fit RMSD	Goodness of fit r^2
		Experimental data			
POU	145.772	245.187	0.184	5.974	.950
POR	52.652	331.963	0.159	8.280	.936
IRP	27.034	1.884		4.202	.706
		Simulation data			
POU	192.656	178.635	0.224	3.619	.974
POR	83.289	294.916	0.209	3.585	.990
IRP	39.506	3.554		4.101	.900

Note. Functions for POU and POR are power functions of the form $y = a + bP^{-c}$, where P = practice. Functions for IRP are linear functions of the form $y = a - bP$, where P = practice. RMSD = root mean squared deviation between observed and predicted values, in milliseconds; POU = performance on unique stimuli; POR = performance on repeating stimuli; IRP = inverted repetition priming.

pretrained on this corpus for four epochs. This pretraining was meant to correspond to a situation where there is preexperimental practice in the task as well as extensive familiarity with certain specific stimuli that occur very frequently. After pretraining the model in this manner, the experimental simulation was run. Instead of using stimuli from the actual digit-entering task experiment (as in the two simulations previously described), we used 40 stimuli in each epoch (block). Thirty of these stimuli were unique to each epoch. The remaining 10 stimuli appeared in every epoch and thus constituted the repeating stimuli. These were the same 10 stimuli that the model had been repeatedly exposed to during pretraining. The model was presented with 24 such experimental epochs.

Simulation results are shown in Figure 15. Figure 15a shows the functions for performance on unique stimuli and performance on repeating stimuli. As can be seen, performance on unique stimuli improves with practice, whereas performance on repeating stimuli exhibits virtually no improvement. Figure 15b shows the derived skill learning and repetition priming functions. As can be seen, skill learning increases with practice, but repetition priming actually decreases with practice. This pattern of results corresponds with Case IV in Figure 7. The simulations thus demonstrate that a pattern in which there is improvement in performance on unique stimuli with practice, but little or no improvement in performance on repeating stimuli, can arise even in a single mechanism; consequently, so can a pattern of results in which skill learning increases while repetition priming decreases. Case IV from Figure 7 is thus fully compatible with the existence of a single underlying mechanism. This concludes our demonstration that all the possible patterns of relationship shown in Figure 7 are indeed consistent with the possibility of a single underlying mechanism.

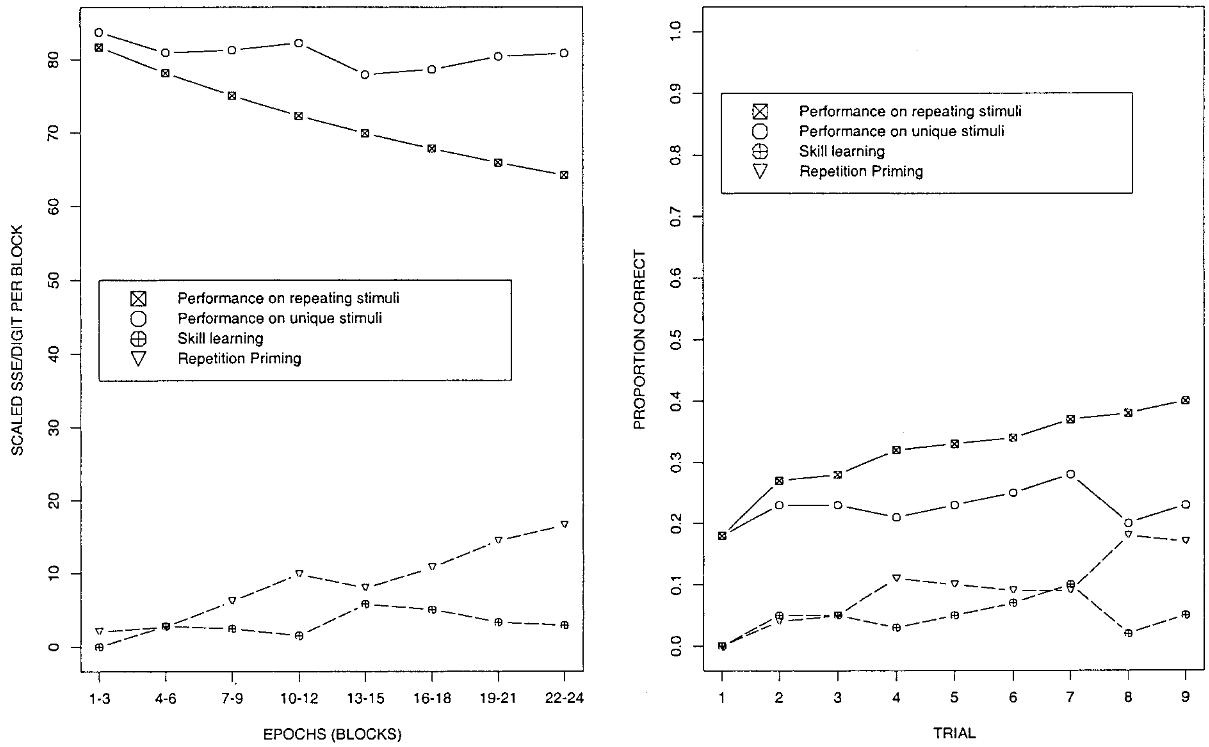
General Discussion

This article has focused on the relationship between skill learning and repetition priming. There has been controversy over whether they reflect a single underlying mechanism (Kirsner & Speelman, 1993, 1996; Logan, 1990; Schwartz & Hashtroudi, 1991), and this question has important repercussions for thinking about the nature of procedural memory as a whole. The analysis in the first part of the article indicated that previous arguments regarding the separation of mechanisms that underlie skill learning and repetition priming are invalid. Specifically, we showed that the lack of a power function for the repetition priming measure does not have any bearing on the processing relationship between the phenomena of skill learning and repetition priming. We showed that the lack of correlations between the skill learning and repetition priming measures also does not bear on the processing relationship between the phenomena. We further showed that none of the possible patterns of increase or decrease between the skill learning and repetition priming measures constitutes evidence that the two phenomena are based on separate mechanisms. These analyses do not of course show that a single mechanism does underlie skill learning and repetition priming. In the second part of this article, we therefore focused on the complementary demonstration that skill learning and repetition priming can indeed arise from a single learning mechanism. We outlined our theory of procedural memory and then presented a computational model incorporating that theory. We showed that the model provides an accurate account of a fairly complex body of behavioral data from the digit-entering task. In particular, the model showed improvement in performance on both unique and repeating stimuli across blocks and exhibited both skill learning and repetition priming. The model also exhibited a number of the apparent dissociations between the skill learning and repetition priming measures that have been taken as evidence for dual mechanisms. All of these effects were exhibited, even though the model consists of only a single learning mechanism, which provides further evidence that the presence of such dissociations in empirical results does not imply the existence of dual mechanisms. The fact that this single-mechanism model provided a good fit to a complex pattern of empirical data suggests that these various demonstrations deserve to be taken seriously.

Necessary and Sufficient Conditions

Our discussion so far may have suggested that skill learning and repetition priming are both necessary consequences of performance in any task. This is not strictly true. In the first place, it may be the case that the transducers (sets of weighted connections) that are involved in performance of a particular task have been tuned to all the transductions in the mapping to such a degree that no further tuning is possible. Under these conditions, practice will yield no further improvement in performance on either repeating or unique stimuli and hence no skill learning or repetition priming. Note that this corresponds to Case IIIb in Figure 7.

Next let us consider the more usual case where there is room for adjustment in the weighted connections. We can identify necessary and sufficient conditions for the occurrence of skill learning and repetition priming. Before doing so, however, we need to make one further observation and introduce some definitions. The observation is that in general, if there is enough room for the weighted connections (i.e., the transducer) to be tuned in response to unique stimuli, then there is necessarily even more room for the transducer to be tuned in response to repeating stimuli. That is, if performance on unique stimuli is improving with practice, then



(a) (b)

Figure 14. Repetition priming without skill learning. (a) In a single mechanism: digit-entering task simulations after extensive pretraining. Repetition priming increases while skill learning exhibits no clear increase. (b) Data from Schwartz and Hashtroudi (1991), Experiment 1: word-fragment completion. Here, too, repetition priming increases while skill learning exhibits no clear increase. The pattern is very similar to that seen in the simulation results.

performance on repeating stimuli must necessarily be better than that on unique stimuli; if performance on unique stimuli is improving, then the performance on repeating stimuli curve must lie below it. Equivalently, if skill learning occurs, there necessarily must be some repetition priming (because the performance on repeating stimuli curve must lie below the performance on unique stimuli curve).

This means that Case Id in Figure 7 is proscribed by our processing theory. That scenario shows improving performance on unique stimuli, but the performance on repeating stimuli curve is no lower than the performance on unique stimuli curve. In our earlier discussion of the scenarios in Figure 7, we were concerned with logical possibility and with showing that none of the logical possibilities provides evidence for separate underlying mechanisms. What we are now claiming is that one of these logically possible scenarios is humanly impossible; our processing theory makes a strong prediction in this regard. If Case Id were to be observed in a multiple-repetition paradigm, we would be forced to revise our processing account. Note, however, that even if Case Id were ever observed, this would still not prove that there are different mechanisms that underlie skill learning and repetition priming. It would show that our present single-mechanism account of the relationship between skill learning and repetition priming is inadequate; it would not show that single-mechanism accounts are in principle inadequate.

This brings us to definitions. We need to distinguish between *systematic mapping* and *arbitrary mapping*. In a systematic mapping, similar stimuli at the first level of representation map onto similar representations at the second level. If stimulus A is similar to stimulus G at the first level of representation, then the transduced representation A' will be similar to G' at the second level of representation. In an arbitrary mapping, there is no such guarantee. Even if A and G are similar at the first level of representation, A' and G' may be quite dissimilar at the second level. To take an example, the mapping between a word's written form and its spoken sound form is in general a systematic mapping, at least to the extent that the spelling system of the language is phonetically consistent. In contrast, the linguistic mapping between a word meaning and its spoken sound form is in general arbitrary; indeed, this is viewed almost as a defining feature of language in general. For example, there is no guarantee that two semantically similar objects such as a mug and a cup will have names that are phonologically related to each other. It should be noted that in most languages, there are partial regularities in the mapping: for exam-

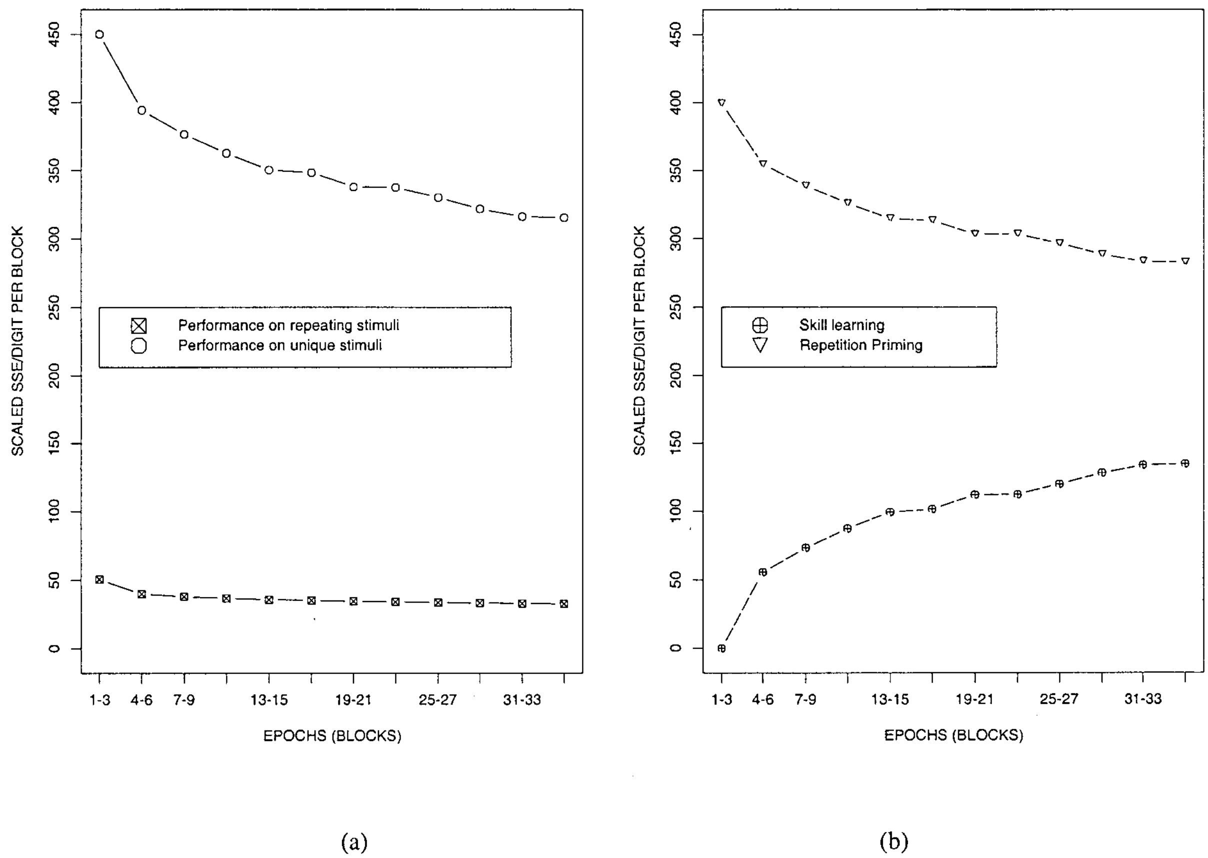


Figure 15. Increasing skill learning with decreasing repetition priming in a single mechanism. The simulation shows improving performance on unique stimuli but relatively constant performance on repeating stimuli. This gives rise to increasing skill learning with decreasing repetition priming. These effects all arise from the single mechanism that is incorporated in the model. (a) Functions for performance on unique stimuli and performance on repeating stimuli. (b) Functions for skill learning and repetition priming.

ple, certain aspects of meaning such as plural number may tend to be expressed consistently in a particular fashion (by adding *-s*, for example, in English). Nevertheless, the mapping is for the most part arbitary.

One further distinction needs to be made. Up to this point, we have assumed implicitly that when stimuli are repeated, they always map onto the same output response.[11] We now relax this assumption. Assume that certain stimuli are presented repeatedly to a given transducer, that is, to a given set of weighted connections. These, of course, constitute what we have termed repeating stimuli. Under our relaxed assumptions, the transduction (and hence the appropriate output response) required for a particular repeating stimulus may be the same on each repetition, or it may be different on different repetitions. We refer to the first case as consistent repetition, and to the latter case as inconsistent repetition of stimuli.

We can now examine the circumstances under which there will be improvement in performance on unique stimuli and performance on repeating stimuli. The possibilities are summarized in Table 4. The first row shows expectations for performance on unique stimuli under conditions of arbitrary mapping and systematic mapping. We would expect no improvement in performance on unique stimuli in an arbitrary mapping. In our formulation, such improvement arises from generalization in a set of weighted connections (which constitute the transducer). If the weights encode (approximate) an arbitrary mapping, there is, by definition, no basis for generalization. Exposure to unique stimuli in a systematic mapping, however, may lead to generalization, if the unique stimuli provide a broad enough sampling of all possible input–output pairings. That is, the necessary and sufficient conditions for improvement in performance on unique stimuli are (a) that the mapping must be systematic and (b) that practice must cover a "sufficient" portion of the set of possible transductions in the mapping. The sufficiency of a particular regime of practice is dependent on the particular mapping and on the particular stimuli

[11] That is, we have assumed that our mappings are functions. A function is a mapping in which each input representation maps onto exactly one output representation. However, our mappings may not always be functions. It may be the case that a given input representation maps onto more than one output representation. To approximate such a mapping, a transducer has to be able to transduce a given input stimulus into different possible outputs.

Table 4
Conditions Under Which Improvement Can Be Expected in Performance on Unique Stimuli and Performance on Repeating Stimuli

Performance on		Arbitrary mapping	Systematic mapping
Unique stimuli		No improvement	Improvement (if sufficient sampling)
Repeating stimuli	Consistent repetition	Improvement	Improvement
	Inconsistent repetition	No improvement	Impossible

sampled. A corollary to this is that there can be improvement in performance on novel stimuli even without a regimen of presentation of "unique" stimuli as we have defined them. If a set of repeating stimuli is presented to a transducer, there may be general improvement in performance on unique stimuli, if the mapping is systematic and if the repeating stimuli sample the mapping broadly enough.

What about the necessary and sufficient conditions for improvement in performance on repeating stimuli? There are potentially four situations to be considered here, determined by whether there is an arbitray or systematic mapping and by whether there is consistent or inconsistent repetition. As shown in Table 4, one of these situations can be ruled out as impossible: It is impossible to have systematic mapping and inconsistent repetition together. This leaves three other situations. We expect performance on repeating stimuli to improve under consistent repetition, irrespective of whether the mapping is systematic or arbitrary. This is because repeated exposure to these stimuli results in tuning of connection weights for these specific stimuli. If the mapping is systematic, and the repeating stimuli provide broad coverage of the mapping, there may be accompanying improvement in performance on unique stimuli. If the mapping is arbitrary, there will be no accompanying improvement in performance on unique stimuli no matter how broadly the repeating stimuli sample the mapping. But in either case, there will be improvement in performance on repeating stimuli. Under inconsistent repetition, however, we do not expect to see any improvement in performance on repeating stimuli. This is because connection weights cannot be adjusted to yield two or more different transductions for input stimuli. Repeated presentation of the stimuli therefore yields no benefit. The necessary and sufficient condition for improvement in performance on repeating stimuli is that there be consistent repetition.

Much of the discussion in this section has focused on specifying conditions under which improvements in performance on unique or repeating stimuli can be expected to occur, and these amount to testable predictions. Our theory makes a particularly strong prediction regarding the impossibility of Case Id in Figure 7. For the reasons noted earlier in this section, empirical observation of such a scenario (in which there was skill learning but no repetition priming in a multiple-repetition paradigm) would require reconsideration of our processing account. Additionally, our account specifies conditions under which various relative patterns of increase–decrease in skill learning and repetition priming can be expected, as in our computational demonstrations of the possibility of Cases II and IV in Figure 7. Case II (in which there is a flat performance on unique stimuli function together with improvement in performance on repeating stimuli) may be expected when subjects have a very high level of preexperimental skill in the task. Case IV (in which there is a flat performance on repeating stimuli function together with improvement in performance on unique stimuli) may be expected when the specific stimuli that repeat have very high preexperimental frequency familiarity. These specifications also amount to theoretical predictions. Our framework further makes a prediction about the magnitude of repetition priming. Repetition priming is measured, of course, by the separation between the performance on unique stimuli and performance on repeating stimuli functions, and according to our theory, the magnitude of this separation will depend on the degree of similarity between the repeating and unique stimuli. If the specific repeating stimuli that are encountered in a task are similar to the specific unique stimuli that are encountered in the task, the separation between the performance on unique stimuli and performance on repeating stimuli functions will be less than if the two sets of stimuli are dissimilar.[12] Overall, there are several predictions that follow from our theoretical framework, which is therefore falsifiable.

Multiple Levels of Skill Learning and Repetition Priming

The results we described from Poldrack et al. (1999, Experiment 1) showed that performance in the digit-entering task was better for the repeating five-digit strings than for the unique five-digit strings; that is, there was repetition priming for five-digit strings. The results also showed, however, that performance in Session 3 was better for rule-following unique stimuli than for non-rule-following unique stimuli. Stated another way, performance was better for those five-digit strings that contained two-digit sequences that had been repeatedly encountered than for those five-digit strings that contained two-digit sequences that had not previously been encountered; that is, there was facilitation for repeating two-digit sequences. These results emphasize that the effects of repetition are seen at a variety of unit sizes. They also highlight the fact that "items" can exist at multiple levels. Repetition priming is often defined as the difference between performance on repeating versus unique items. In the digit-entering task, repetition priming occurred for five-digit items, as indicated by the superior performance on repeating versus unique five-digit strings. However, the results suggest that two-digit sequences are also items: There is repetition priming for these units also. Thus how an "item" is defined in such tasks needs to be approached with some caution.

We have characterized the phenomena of skill learning and repetition priming as being manifestations of the same underlying process of incremental adjustment in the transductions on which a particular task depends. A corollary of this is that skill learning and repetition priming can arise in multiple transducers and thus at multiple levels in a system. Let us further explore this notion, applying it to understanding the differences between the multiple-

[12] We thank Stellan Ohlsson for pointing this out (personal communication, February 1998).

repetition and study–test repetition priming paradigms. Consider the hypothetical task that was introduced earlier in which words are presented visually, and a participant has to say how many syllables each word contains. Performance in the task is measured by response latency. In a multiple-repetition paradigm, there would be several blocks of practice. Some of the words would appear in every block (the repeating stimuli), while other words would appear exactly once (the unique stimuli). That is, each block would consist of repeating stimuli as well as unique stimuli. We would expect an improvement, over blocks, in task performance on unique stimuli and on repeating stimuli, and we would expect performance on repeating stimuli to be better than performance on unique stimuli. That is, we would expect to see both skill learning and repetition priming.

The study–test version of the same task would consist of two phases. Prior to performing the actual syllable-judgment task (referred to as the "test" phase), there would be a "study" phase. The study phase would include exposure to some of the stimuli that would later appear in the test phase. The purpose of the study phase would be to provide prior familiarization with these stimuli, but not task-specific facilitation. For example, at study, participants might be asked to rate words for correctness of spelling. After a distractor-filled interval, the test phase would begin. Now, participants would perform the syllable-judgment task. Some of the stimuli would be drawn from the previously studied words. In this paradigm also, we would expect to see superior performance on the repeated (studied) stimuli than on the nonrepeated stimuli. That is, we would expect to see repetition priming.

Conceptualizing the task as a system of transductions helps us to see that the locus of priming is different in the two cases. Figure 16 is a schematic of the transducers that are involved in the syllable-judgment task. The first transducer (A) enables the sensory stimulus to be converted into an internal representation of the word through processes of word identification. For the present analysis, we do not need to specify what the representations or processes are in this transduction. It is enough to point out that this transducer must exist, quite independent of whether a participant has ever performed a syllable-judgment task, because this is the transducer that enables single words to be read. Furthermore, assuming that the participant is a normal adult reader, this transduction is highly overlearned. The second transduction shown (B) is task specific. Its input is the representation delivered as the output of the first transduction; this representation must be further transduced into information about the number of syllables that are contained in the represented word. This transduction is much less practiced than is the first one. However, performance of the syllable-judgment task relies on both of the transducers.

Now let us consider performance improvement in the multiple-repetition paradigm. Transducer A (which consists of the ability to read words) is overlearned. This means that there is little room for general improvement in this transducer: Reading the relatively small number of words that are contained in the experiment will lead to little, if any, general improvement. However, we would expect a significant improvement in performance on unique stimuli in the second, more task-specific transducer B, because this transduction is relatively unpracticed. Thus, the locus of improvement in performance on unique stimuli will be primarily in transducer B. What about the locus of improvement in performance on repeating stimuli? We would expect substantial adjustment toward

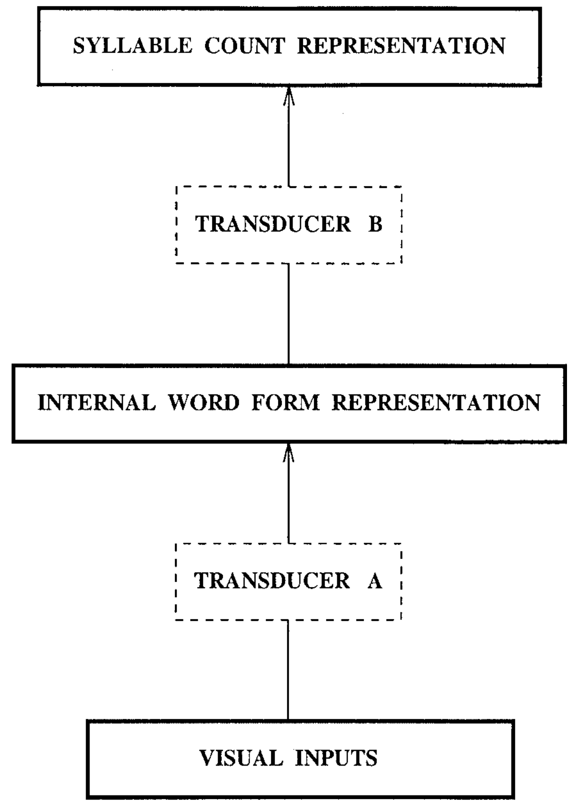


Figure 16. Levels of representation and transductions in a hypothetical syllable-counting task.

repeating stimuli in the connection weights in transducer B, and hence substantial improvement in performance on repeating stimuli. In addition, we would expect some improvement in performance on repeating stimuli in transducer A as well: Repeated exposure will lead to some facilitation for these specific stimuli, although not to generalized facilitation. To summarize, the locus of improvement in performance on unique stimuli is largely or entirely in the task-specific second transduction B; accordingly, the locus of skill learning is at this second level. The locus of improvement in performance on repeating stimuli is in the second transduction as well as in the first one; accordingly, repetition priming occurs in both of the transductions, that is, at multiple levels.

Let us next consider performance improvement in the study–test paradigm. At study, certain stimuli are read, that is, transduced at the first level. Incremental weight adjustments occur for these stimuli. At test, some of these stimuli are read again. There will be an advantage in processing of these stimuli by transducer A, relative to those test stimuli that were not studied. That is, adjustment in connection weights in transducer A will lead to repetition priming. As in the previous case, however, the relatively small number of words contained in the experiment is unlikely to lead to generalized improvement in transducer A, and so we would expect no skill learning in transducer A. Turning to the task-specific transducer B, we can note that it is only exposed to any stimulus once. That is, from the point of view of this transducer, there are no repeating stimuli, only unique stimuli. Therefore, there can be no repetition priming in this second transducer; we would, however, expect skill learning to occur in transducer B. Comparing the multiple-repetition and study–test versions of the task, we see that in both versions, the locus of skill learning is in transducer B. However, the locus of repetition priming differs in the two variants of the task: In the multiple-repetition paradigm, it is in both transducers, whereas in the study–test paradigm, repetition priming can only occur in transducer A.

It is important to distinguish our formulation from a view in which, in any given task domain, there are two different modules (representing skill learning and repetition priming) that process different kinds of information in that domain, with one module processing information that is relevant to skill learning and the other module processing information that is relevant to repetition priming, but with both modules operating via an incremental tuning process of the kind we have described. Our formulation is crucially different in the following way. In our framework, the modules (i.e., processors or transducers) that are deployed in a task indeed operate on different representations and thus process different kinds of information (e.g., transducers A and B in the syllable-judgment task). However, the distinction between the transducers and the distinction between the kinds of information they process corresponds not to a distinction between skill learning and repetition priming, but to a distinction between functional components or processing stages of the task, or both. Even if the locus of repetition priming is in transducer A and the locus of skill learning is in transducer B, as in the study–test version of the syllable-judgment task, this is not because transducer A is specialized for repetition priming and transducer B specialized for skill learning; rather, the differing loci of skill learning and repetition priming emerge from the operation of transducers that map onto different task components.

The difference between study–test paradigms and multiple-repetition paradigms is a significant one. In a multiple-repetition paradigm, the task remains the same throughout, and this means that the same transducers are exposed to both the repeated and nonrepeated stimuli. Even so, the loci of any observed skill learning and repetition priming effects may be in different transducers, as just discussed. In a study–test paradigm, there is usually a change in task between study and test, and the transducers that are exposed to the repeated and nonrepeated stimuli will therefore usually be different. The loci of any observed skill learning and repetition priming effects are therefore even more likely to be in different transducers than is the case in a multiple-repetition paradigm, as also just discussed. Conclusions about the mechanisms that underlie observed patterns of skill learning and repetition priming in a study–test paradigm must therefore be approached with even more caution than in a multiple-repetition paradigm. For example, Schwartz and Hashtroudi (1991, Experiments 1 and 2) studied repetition priming and skill learning in a partial-word identification task, using both a multiple-repetition paradigm (Experiment 1, which we have already discussed in this article) and a study–test paradigm (Experiment 2). In the study–test paradigm, they found that repetition priming was no higher in a group of participants who were skilled in the task than in participants who were not skilled in the task and interpreted this result as a dissociation between repetition priming and skill learning. However, as we discuss in detail in the Appendix, this interpretation does not hold up when a careful consideration is made of the nature of the study–test paradigm that was used.

In an earlier section, we pointed out that differences that have been observed between repetition priming and skill learning in neuropsychological studies have often been based on different tasks (see the discussion at the end of the Theoretical Analysis section). Because different tasks involve different sets of transducers, interpretation of these results as indicating a dissociation between repetition priming and skill learning is invalid. A similar point arises in interpreting the results of neuropsychological studies that use a study–test paradigm. To see this, let us return to the study–test version of the syllable-judgment task that was described above. Let us suppose that an individual who will participate in this task has a lesion that selectively impairs transducer A (see the earlier discussion of this task), thus reducing the facilitation (adjustment of connection weights) that can arise in this transducer. When the individual undertakes the study phase, the lesioned transducer A supports a reasonable level of performance in the task (which is simply to read each word), but only a small amount of facilitation toward these stimuli takes place. At test, therefore, in the syllable-judgment task proper, there is little or no benefit in transducer A for the previously presented words, that is, little or no repetition priming in transducer A. We would not expect skill learning either in transducer A, both because this is a highly overlearned mapping and because facilitation is impaired in this transducer. Turning to transducer B, at test, there cannot be any repetition priming, because, as discussed earlier in this section, this transducer did not previously process the studied stimuli. However, over the course of the test phase trials, there might be skill learning in transducer B. Thus, in performance of the study–test task overall, we might see no repetition priming but some skill learning.

Such a result might appear to indicate that there had been selective impairment of repetition priming but not of skill learning in what appeared to be a single task, and thus that repetition priming and skill learning are dissociable processes. However, such a conclusion would fail to take into account that the transducers were different at study and test, and that the impairment was selective to processing that is differentially involved in the study and test phases, rather than selective to repetition priming. Note also that such a result would not refute our earlier prediction (see the *Necessary and Sufficient Conditions* subsection) that repetition priming without skill learning is impossible in a multiple-repetition paradigm or, more generally, in a task in which the same transducers are involved at all points. The hypothetical result currently under discussion is permissible in our framework because in the syllable-judgment study–test task, the transducers that are involved are not constant across study and test.

As a further example of the relevance of such task analysis to neuropsychological investigation, let us consider the case of patient M.S., whose pattern of priming following a right occipital lesion has been examined in a number of studies (e.g., Fleischman et al., 1995; Gabrieli, Fleischman, Keane, Reminger, & Morrell, 1995; Vaidya, Gabrieli, Verfaellie, Fleischman, & Askari, 1998). In the most pertinent study, M.S. and various types of controls participated in a study–test paradigm (Gabrieli et al., 1995). In the study phase of the task, each participant read aloud 24 visually presented words. In the test phase, which was a perceptual identification task, 48 words (24 from the study phase and 24 unstudied words) were shown very briefly to each participant, whose task was to identify the word. Of interest was how much better participants would be at identifying studied versus unstudied words in the test phase, the measure being the mean duration needed to identify the two types of words. The controls had significantly shorter durations to identify studied than to identify unstudied words, thus exhibiting priming. However, M.S. did not exhibit priming for the studied words.

How should we interpret these results? Let us begin by considering what transducers might have been involved in the task at study and at test. The processing system engaged at study can be broadly conceptualized as consisting of a transducer A that maps from the visual representation of a word to an internal word form representation (essentially identical to transducer A in the syllable-judgment task) and a transducer B that maps an internal word form representation to a sequence of articulatory gestures (i.e., pronouncing the word out loud). The perceptual identification task at test involves speeded presentation of stimuli, and, presumably, speeded responses. Let us make the simplifying assumption, however, that the system engaged at test consists of exactly the same transducers A and B that were involved in the study phase. The priming that was observed in controls and absent in patient M.S. was presumably due to the presence and absence, respectively, of facilitation in transducer A (as there is no reason to believe that M.S. had any impairment in transducer B). We also note that, in reality, transducer A is likely to be a complex system that consists of several subtransducers that collectively implement the mapping from visual word form presentation to internal word form representation; the impairment to transducer A may be localized to some subset of its subtransducers. Let us now return to interpretation of the results.

One interpretation would be that the presence of priming in controls and its absence in M.S. indicate the operation of a specialized visual implicit memory system (e.g., Gabrieli et al., 1995). This view amounts to hypothesizing that transducer A (or some subtransducer within A) is specialized for visual implicit memory and is selectively impaired. An alternative view that we favor is as follows. In an unimpaired system, presentation of a stimulus at study leads to facilitation in some subset of A's subtransducers. What facilitation is revealed at test will depend on what aspects of stimulus presentation are manipulated at test. Impairment in some subset of subtransducers within A will affect the operation of those subtransducers, including the operation of procedural learning and hence facilitation within them during study. At test, the facilitation that is obtained in the overall task may differ from what would have been obtained in an unimpaired system, depending on the specific loci of impairment within A and on which aspects of the stimulus are repeated from study to test. However, neither transducer A nor its subtransducers are viewed as being specialized implicit memory systems; rather, transducer A is a complex visual stimulus processing system that incorporates procedural learning (and hence implicit memory) as a fundamental part of its operation in every subtransducer.[13] On this view, M.S. has an impairment somewhere in transducer A, and this impairs the operation of that subcomponent of A, including procedural learning in that subcomponent; this leads to reduced facilitation in that subcomponent of A at study, which is revealed at test. It should be possible to distinguish between the two views that we have outlined, because the second view makes the testable prediction that if patient M.S. manifested an impairment of priming in a multiple-repetition version of the perceptual identification task, he would also manifest an impairment in skill learning in the task; this follows by *modus tollens* from our earlier general prediction that if skill learning occurs in a multiple-repetition paradigm, it must necessarily be accompanied by repetition priming.

In summary, skill learning and repetition priming can occur at multiple levels within a system. Furthermore, the locus of repetition priming may be rather different in different versions of what prima facie might appear to be the same task. These considerations suggest that an understanding of repetition priming and skill learning effects that are obtained in any given task needs to be preceded (a) by a task analysis that conceptualizes the various transductions involved in the task and (b) by an analysis of the loci of improvements in performance on unique stimuli and performance on repeating stimuli. The present framework encourages careful analysis of the mappings and transducers that are involved in a task and highlights the importance of examining whether the effects observed in two tasks (or in variants of what appears to be the same task) are in fact comparable.

Broader Implications

Skill learning, repetition priming, and automaticity. In this article, we have sketched a theory of procedural memory, focusing chiefly on its implications for skill learning and repetition priming. How does this theory relate to broader issues in memory and cognition? We will begin by considering the phenomenon of *automaticity*, which is an important subject in the study of skill acquisition. Automaticity refers to the mode of cognitive processing in which certain activities can be performed quickly, effortlessly, with little awareness, and with little need for conscious thought (Logan, 1988). According to the present theoretical framework, automatization in a task arises as a result of the same processes of incremental tuning that we have described as underlying improvements in performance on unique stimuli and performance on repeating stimuli. On our account, automaticity is a state in which the transductions for certain stimuli have been tuned through repeated exposure, to such a degree that performance on those particular stimuli has become highly facilitated. When task performance is examined with respect to these particular stimuli, performance gives indication of having become "automatized."

We wish to highlight three aspects of this formulation. First, automaticity is merely a state in which performance on repeating stimuli has improved greatly; it is the near-asymptotic part of the performance on repeating stimuli function. Performance on nonrepeating stimuli will necessarily be inferior, because the connection weights have not undergone repeated tuning toward these stimuli. The item-specific nature of automaticity therefore falls out of this conceptualization, as it does out of Logan's (1988) instance

[13] Similarly, in this view, dissociations between perceptual priming and conceptual priming in healthy people (e.g., Blaxton, 1989; Roediger & Blaxton, 1987) do not reflect distinct implicit memory systems but rather the differential involvement of different transducers, each of which incorporates procedural learning and can thus exhibit priming, which therefore appears as either more perceptual or more conceptual. We return to these points in our discussion of multiple memory systems and components of processing.

theory of automaticity. The second point we wish to emphasize is that whether performance is judged to be automatic depends on what criterion of performance is used, as Logan has previously noted. The term automaticity can thus be seen as being merely a label that is applied to performance on repeating stimuli when it improves beyond some arbitrarily defined criterion. Third, given that in our framework, automaticity is merely the extremum of the phenomenon of repetition priming, we can specify the conditions under which automaticity will arise, as we have already specified these conditions for improvement in performance on repeating stimuli. Referring back to Table 4, we can see that performance on repeating stimuli will improve whenever there is consistent repetition of stimuli (provided that there is some room for adjustment of connection weights). That is, the necessary and sufficient conditions for automaticity are (a) consistent repetition and (b) plenty of it. This theoretical prediction is borne out by empirical evidence, which indicates that automaticity is acquired only when stimuli are mapped consistently onto the same responses throughout practice (Logan, 1978, 1979; Schneider & Fisk, 1982; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Our framework offers an account of this finding.

As Logan (1988) has noted, the term automaticity has also been used in another sense: to refer to the kinds of effects observed in Stroop tasks (Stroop, 1935). Cohen, Dunbar, and McClelland (1990) presented a computational model that accounted for a wide variety of phenomena in Stroop paradigms. In their account, performance in tasks such as word reading and color naming lies along a continuum of strength of processing. When a task is highly practiced, the strength of processing in the pathways that enable its performance is high. Stroop-like interference and facilitation effects arise from the relative strength of processing in the pathways in competing tasks. The starting point in the work of Cohen et al. (1990) was Stroop-like paradigms. The starting point for the present work was the kinds of paradigms in which repetition priming has been studied. Both approaches converge on the notion that the processing factor that leads to automaticity is the effectiveness of the underlying processing pathways, and that this effectiveness is altered by incremental tuning of connection weights. The present work thus complements that of Cohen et al. in highlighting the relationship between two senses of the term *automaticity*.

Multiple memory systems and components of processing. Our framework also makes contact with another significant issue in the study of memory, namely, the debate over whether memory is better characterized as being a single system or as consisting of multiple systems. This debate was recently reviewed in an insightful article by Roediger et al. (1999). As Roediger et al. note, advocates of multiple memory systems have proposed that there are distinct memory systems for such things as episodic versus semantic memory (Tulving, 1972) and procedural versus declarative memory (Cohen & Squire, 1980; Cohen & Eichenbaum, 1993) as well as a presemantic perceptual representation system (PRS), which is a memory system that is distinct from the other systems and that itself has a number of distinct memory subsystems such as a word form PRS and a structural description PRS (Tulving & Schacter, 1990). Advocates of a unitary view of memory, on the other hand, have argued that a focus on processing operations is critical to understanding of memory (e.g., Craik & Lockhart, 1972; Kolers, 1975) and have suggested that the distinction between data-driven and conceptually driven processing needs to be considered in interpreting results from different memory tasks that have been thought to reflect the operation of different memory systems (e.g., Blaxton, 1989; Jacoby, 1983).

Roediger et al. (1999) argue that each point of view has its strengths and weaknesses, and that a more illuminating way of thinking about memory is one that offers a resolution of the two viewpoints. In particular, as in earlier work (Roediger & McDermott, 1993), they endorse the "components of processing" approach put forward by Moscovitch (1992, 1994). This approach is based on the assumption that performance of each task requires the operation of many components, some of which are common to tasks and some of which are not. In this view, any two tasks that are dissociated must differ in at least one task component.

Roediger et al. (1999) illustrate the application of such an approach by means of a thought experiment with three groups of participants. In this thought experiment, each group is given a series of word stems (the first three letters of words, e.g., COU ______) as a cue for their task, so that the perceptual display is the same for all groups. One group is instructed to say the first word that comes to mind that fits the stem. A second group performs the same task, with the same instructions and the same stems, but this group had previously studied a list of words. Some of the stems they now see represent words that appeared on the previous list (e.g., COURAGE). The third group of participants studies the same list of words as that studied by the second group and receives the same word stems at test. However, at test, they are instructed to use each stem to think back to an item in the studied list and to produce words to fit the stems only if the word had occurred in the list. As Roediger et al. (1999) note, the procedure for the second group instantiates a typical implicit memory test, whereas the procedure for the third group instantiates a typical explicit memory test. The first group represents a baseline against which repetition priming is calculated in the implicit memory test.

Roediger et al. (1999) note that from the point of view of Moscovitch's (1992, 1994) components of processing theory, the tasks would be analyzed as follows. The first condition, completing word stems, relies on a complex set of processing components. This set of components represents a cognitive system that achieves the task of producing words from three letters. When participants have studied some of the words prior to completing stems, as in the second condition, components of the system responsible for completing the word stem might change—some might be added, and others might drop out. But these are changes in components of the system rather than the engagement of an independent system. For the third group of participants, the system is once again different, with different components that are involved in remembering being brought into play. The differences in components between the three tasks could be labeled as "memory systems," but in the components of processing approach, this is not regarded as particularly fruitful.

The theoretical framework we have presented here is completely consonant with Roediger et al.'s (1999) articulation of the components of processing view. At several points in the present article, we have outlined our view that a particular task is carried out by a particular configuration of transducers, and that this ensemble constitutes the system that underlies performance of that task. Furthermore, we completely agree that when two different tasks deploy different components (i.e., different, albeit possibly overlapping, sets of transducers), these differences do not necessarily mean that there are independent or different "memory systems" involved.

It may be worth elaborating on this discussion to clarify our own views about the multiplicity (or otherwise) of memory systems. In our view, there is a fundamental distinction between two kinds of biological learning and memory mechanisms, which we have termed procedural memory and declarative memory, respectively (Cohen & Squire, 1980; Cohen & Eichenbaum, 1993). It should be noted that there is now considerable evidence regarding the neural basis of such a dissociation between two kinds of memory systems, as acknowledged by Roediger et al. (1999). The declarative memory system is believed to be subserved by the hippocampus and related medial temporal lobe structures (we will refer to this loosely as "the hippocampal system"); these structures provide for the inital encoding of memories that involve arbitrary conjunctions and also for their eventual consolidation and storage in neocortex (e.g., Cohen & Squire, 1980; Mishkin, Malamut, & Bachevalier, 1984; Squire, Knowlton, & Musen, 1993). The procedural memory system, which provides for the learning and processing of motor, perceptual, and cognitive skills, is believed to be subserved by learning that occurs in nonhippocampal structures such as neocortex and the basal ganglia (e.g., Cohen & Squire, 1980; McClelland et al., 1995; Mishkin et al., 1984; Squire et al., 1993). It has further been suggested that there is a computational reason for the existence of two memory systems (McClelland et al., 1995). According to this view, the neocortex processes distributed representations. Such representations have the desirable property of providing a basis for similarity-based generalization. However, in a computational system that uses distributed representations, the learning of arbitrary associations must be slow and interleaved (McClelland et al., 1995), if the system is not to suffer from catastrophic interference, that is, the overwriting of previously existing knowledge (McCloskey & Cohen, 1989). However, the encoding of the kinds of arbitrary associations that constitute declarative memory (episodes and new facts, which are arbitrary conjunctions of items of information) must often occur swiftly, within a single encounter. It therefore cannot proceed via slow cortical learning. It has been proposed that the role of the hippocampal system is to convert distributed representations into localist nonoverlapping ones and to establish fast mappings between such converted representations. That is, the hippocampal system performs fast learning, based on orthogonalized representations, and thus provides a basis for the swift encoding of arbitrary associations of the kind that constitute episodic and factual information (Cohen & Eichenbaum, 1993; McClelland et al., 1995). Neocortex and the hippocampal system thus perform complementary learning functions and provide the basis for procedural and declarative memory, respectively. The neocortical learning system proposed by McClelland et al. (1995) is essentially equivalent to the present view of procedural memory learning.

In designating the kind of learning that occurs in nonhippocampal structures as procedural learning, we adopt the view that neocortical and other nonhippocampal structures can usefully be characterized as operating according to a common learning mechanism that stands in contrast to hippocampal learning mechanisms and takes the form of slow, associative, error-driven learning. The procedural memory system is merely the set of structures that can be characterized as operating via this procedural learning mechanism. Clearly, there are a host of differences in the morphology of different nonhippocampal structures, but, ex hypothesi, these structures can nevertheless be usefully characterized as operating according to a common learning principle. It is worth noting, furthermore, that our assumption of an error-driven learning process is consistent with the widely adopted Rescorla-Wagner model of conditioning (Rescorla & Wagner, 1972). Variants of this basic model have been proposed as accounts of a variety of conditioning phenomena relying on neural structures including neocortex and cerebellum (see Bartha & Thompson, 1995, and Schmajuk, 1995 for review). As pointed out by Sutton and Barto (1981), the learning principle that underlies the Rescorla-Wagner model is equivalent to the error-driven Widrow-Hoff learning algorithm (Widrow & Hoff, 1960). Important for present purposes, this learning algorithm (often called the *delta rule*; e.g., Rumelhart et al., 1986), typifies the kind of error-driven learning we assume for procedural learning. Our assumptions about the nature of procedural learning therefore make contact with well-established traditions of inquiry.

As a final point, it may be noted that we use the terms procedural learning and procedural memory almost interchangeably. This is because procedural learning occurs in the connections between cortical (or other nonhippocampal) representations, and procedural knowledge or memory inheres in the same connections and representations. The situation is slightly different for declarative knowledge. Acquiring new declarative knowledge depends on the hippocampal system. Eventually, however, this knowledge will be redistributed to cortical tissue via a process of consolidation (provided that the hippocampal system is intact), and this knowledge will then reside in cortical tissue in the form of connections between cortical representations (Cohen & Eichenbaum, 1993; McClelland et al., 1995). Thus, at any point in time, older declarative memories will be less dependent on the hippocampal system (because they will have been consolidated into neocortex to a greater extent), whereas more recently formed declarative memories will be more dependent on the hippocampal system. Thus the key mechanisms of declarative learning reside in the hippocampal system, but declarative memory or knowledge is distributed across the hippocampal system and cortex.

To return to the preceding discussion, performance of a given cognitive task may invoke either or both of the procedural and declarative learning mechanisms. Different tasks will require different relative contributions from these two mechanisms. Some tasks (such as explicitly recalling whether a given word was on a previously viewed list) depend crucially on the mechanisms of declarative learning and memory. Other cognitive tasks (such as comprehending language) do not appear to depend crucially on declarative learning and thus appear to be more procedural in nature (although aspects of language acquisition are likely to have a strong reliance on declarative learning, a point to which we will return).

Let us relate this to the idea that there is a particular configuration of processors that underlies performance of any given cognitive task. What it means for a task to rely critically on declarative learning is that the set of processors that is necessary for performance of the task includes structures in the hippocampal system. Cognitive tasks that are used in investigation of implicit memory in general, and of skill learning and repetition priming in particular, are tasks that do not depend critically on the hippocampal system. This must be the case, because these are tasks on which patients with hippocampal amnesia demonstrate normal performance. Therefore, the processors deployed in these tasks must consist primarily of neocortical or other nonhippocampal structures, and the procedural learning that occurs in these tasks occurs in these nonhippocampal processors.

The issue with which we have been concerned in this article is whether dissociations observed between skill learning and repetition priming in this "implicit" class of tasks necessarily implicate different learning procedures within the underlying processors. We have attempted to show that such dissociations do not necessitate the assumption of different learning procedures or mechanisms within the procedural system. Furthermore, we have shown that dissociations of the kind observed between skill learning and repetition priming can be accounted for under the hypothesis that there is a single procedural learning mechanism in operation.

Therefore, on the one hand, there is in our view a real distinction between two different kinds of learning mechanisms, which we term declarative learning and procedural learning, and this view is supported by a great deal of converging evidence. In this respect, we would argue there there are indeed multiple (or rather, dual) memory systems. On the other hand, we would also argue that to interpret differences in performance between tasks, it is critical to analyze the processing components that underlie those tasks, and in this respect we agree with the need for a processing emphasis in the study of memory. We believe there is an important methodological difference between the inferring of a distinction between procedural and declarative memory mechanisms, which is supported by evidence independent of behavioral dissociations between tasks (such as the computational considerations and biological evidence reviewed above) and inferring the existence of memory systems purely on the basis of behavioral dissociations between tasks (which we, along with Roediger et al. (1999), regard as unfruitful).

Overall, our approach is consistent with Roediger et al.'s (1999) formulation of the components of processing view, which emphasizes that (a) any particular task depends on a complex configuration of processors, (b) the deployment of different sets of processors in different tasks does not in and of itself mean that there are different memory systems underlying those tasks, and (c) dissociations between tasks do not in and of themselves mean that there are different memory systems underlying those tasks. However, tasks do differ in the extent to which their underlying configuration of processors depends on procedural or declarative learning mechanisms.

The processing framework we outline in this article thus bears on the debate between multiple systems and processing systems approaches to memory and contributes toward resolution of that debate. In so doing, it is in closer agreement with the "components of processing" approach championed by Moscovitch and colleagues and Roediger and colleagues than might previously have been apparent. In recent years, other authors have also outlined somewhat similar views (Blaxton, 1995; Gabrieli, 1995; Jacoby & Kelley, 1992; Kirsner, 1998).

Procedural memory, declarative memory, and language. According to a view of language learning that has been increasingly adopted in recent years, many aspects of language learning can be understood as proceeding via the gradual, experience-driven adjustment of connection weights between levels of distributed representation. This view has been applied to aspects of language as varied as sentence comprehension (McClelland, St. John, & Taraban, 1989), inflectional morphology (Gupta & MacWhinney, 1992; Hoeffner, 1992; MacWhinney & Leinbach, 1991; Rumelhart & McClelland, 1987), phonology (Dell, Juliano, & Govindjee, 1993; Gupta & Touretzky, 1994), and reading (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989). This is, of course, the view of learning and cognitive processing embodied in the PDP framework (McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986). It is also, however, the view of procedural learning that we have outlined here. In fact, we have hypothesized elsewhere (Gupta & Dell, 1999) that those aspects of language learning that can be characterized as gradual, experience-driven tuning are in fact forms of procedural memory, a suggestion that has also been made by a number of other researchers (e.g., Cleeremans, 1993; Reber, 1993). This suggestion extends the domain of procedural memory to encompass many of the systematic aspects of language learning.

Let us consider a specific example of how ideas about skill learning and repetition priming are applicable to language. We start by noting that a well-known model of errors in speech production (Dell et al., 1993) incorporates essentially the same computational architecture that we have used in our model of the digit-entering task. Like the present model, the Dell et al. (1993) model takes as its input a static representation of an entire sequence (in that model, a word) and produces as its output the elements of the sequence (in that model, phonemes). That is, the model "spells out" its input as a sequential output, just as in our model of the digit-entering task. The model uses the same learning algorithm we use in the present simulations. Thus there is a striking similarity between our formulation of the mechanism underlying skill learning and repetition priming in a digit-entering task and Dell et al.'s (1993) formulation of learning to produce speech sounds.

Furthermore, although Dell et al. (1993) did not focus on this, we can consider the effects of practice on such a model of speech errors. Let us suppose that the model is exposed to a number of word forms in each period of time. In each period of time (such as a day), some of these word forms are ones that have previously been encountered, and some of the word forms are novel, as depicted in Figure 17a. Let us further suppose that the model's task is simply to spell out each input word form representation as a sequence of output phonemes or phonological features. How might the system's basic ability to perform this task develop? Figure 17b shows the pattern of results that we might expect. The ability to spell out familiar forms would improve greatly, as a result of multiple exposures to these forms; this is depicted by the lower curve in Figure 17b. Additionally, we would expect cumulative practice to result in generalized performance improvement, as a result of continual adjustment of connection weights in the system. That is, we would expect the model to become more effective at the process of spelling out word forms in general, irrespective of whether they were familiar. This is shown in the upper curve in Figure 17b.

It should be clear that these effects of repetition on novel and familiar stimuli are precisely those that are studied under the general heading of "skill learning" and "repetition priming" in implicit memory research, as emphasized in Figure 17c and 17d. Our framework thus suggests that the phenomena of skill learning

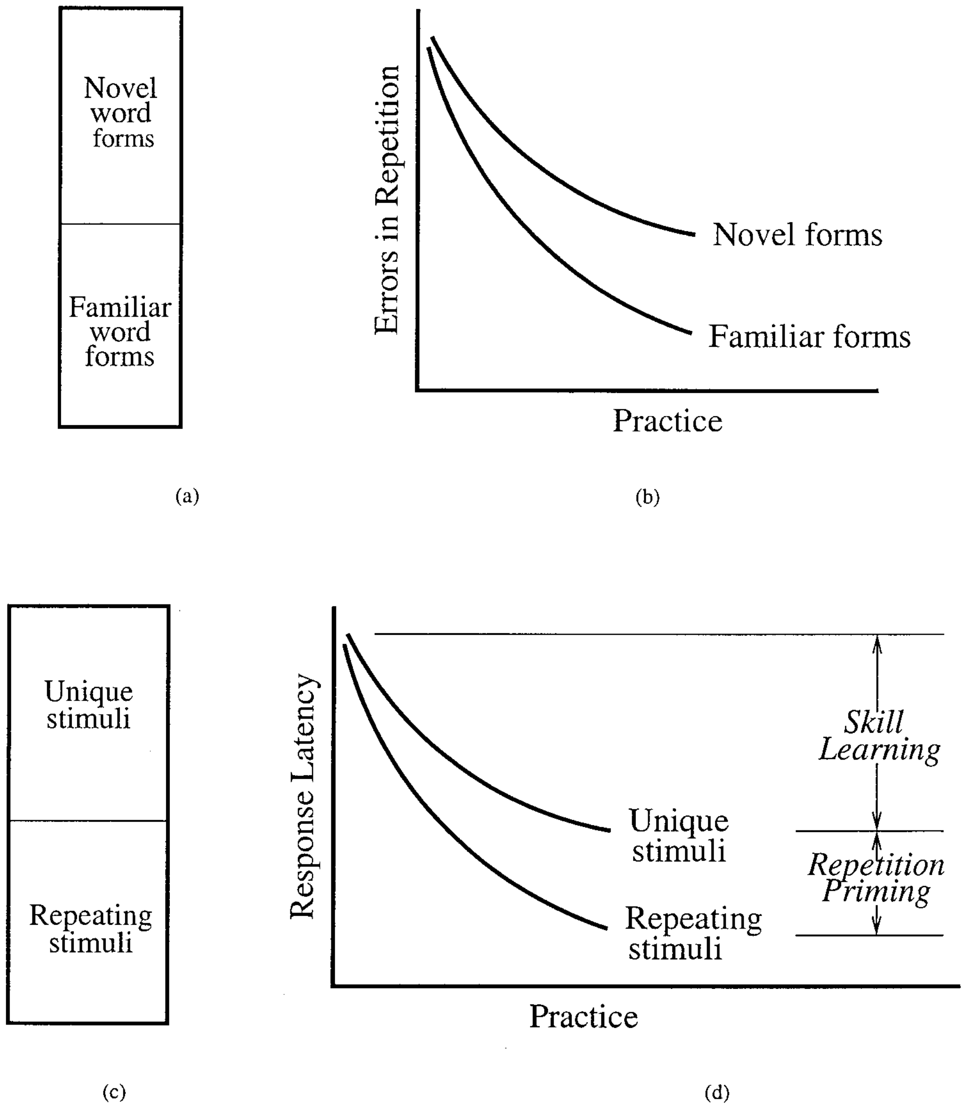


Figure 17. Comparison of expected developmental course of repetition of word forms with skill learning and repetition priming. (a) Hypothetical regimen of environmental exposure to word forms. (b) Expected schedule of development. (c) Regimen of experimental practice in a typical implicit memory multiple-repetition paradigm. (d) Measures of interest in such a paradigm.

and repetition priming are highly relevant to understanding language processing and learning. In fact, we can go further than this. The preceding discussion indicates that skill learning in the language production task is manifested as improved ability to pronounce nonwords (i.e., novel word forms), whereas repetition priming is manifested in improved knowledge of known words. In the study of language, these effects are often viewed as separate phenomena—the learning of phonological rules versus the learning of a lexicon. But to the extent that these are merely skill learning and repetition priming effects within a single speech production system, and given the arguments we have advanced in this article about the mechanisms that underlie skill learning and repetition priming, it can be seen that the linguistic distinction between the "lexicon" and "phonological rules" may be less clear-cut than is often supposed.

As another example of the applicability of ideas about procedural memory to language processing, we can consider the phenomenon known as *structural priming,* which refers to the tendency for speakers to repeat syntactic structures in successive utterances (e.g., Bock, 1986; Levelt & Kelter, 1982). The phenomenon can be demonstrated in experiments in which participants are presented with a mixture of pictures and auditory sentences and are required to repeat each sentence and to describe each picture. If the auditorily presented sentences incorporate a particular syntactic structure (such as the passive construction), then participants become biased to use that structure in their description of the pictures (Bock, 1986; Bock & Loebell, 1990; Bock, Loebell, & Morey, 1992). It has been specifically proposed that structural priming is a form of implicit learning (Bock et al., 1992; Chang, Griffin, Dell, & Bock, 1997; Dell, Chang, & Griffin, 1999). Furthermore, the viability of such a view has been demonstrated in the form of a computational model in which connection weights are adjusted slightly each time a sentence is presented to the model. This incremental tuning process (identical to our view of procedural learning) leads to structural priming in the model (Chang et al., 1997; Dell et al., 1999). The view of procedural learning that we have outlined here thus has relevance to the linguistic domain of syntactic processing as well.

The theoretical framework within which we view procedural and declarative memory also has further implications for aspects of lexical learning. In particular, consideration of the processes that are involved in learning a novel word suggests that learning a new word has two components, and that these components rely differentially on procedural and declarative memory (see Gupta & Dell, 1999, for more detailed discussion). To see this, note that the mappings between sequences of sounds and the internal phonological representation of a word form are systematic mappings, in that similar sequences of input phonemes map onto similar internal word form representations, and similar internal word form representations map onto similar sequences of output phonemes. However, the mapping between word forms and meanings is arbitrary, in that similar internal word form representations are not guaranteed to map onto similar meanings. The import of these observations becomes clearer when we consider that learning a new word can in general be a fast process, occurring within just one or two exposures (Carey, 1978; Dollaghan, 1985). This implies that both the arbitrary semantics and the systematic phonology of a novel word can be learned very rapidly. As noted earlier, the rapid learning of arbitrary mappings is precisely the function that is ascribed to declarative memory (Cohen & Eichenbaum, 1993; McClelland et al., 1995). We therefore might expect that learning the meanings of new words relies heavily on declarative memory, but that procedural memory suffices for learning the phonology of new words. Two predictions follow from this. First, under conditions of impairment in declarative memory, it should be difficult to learn new word meanings quickly. Second, even if declarative memory is impaired, it should nevertheless be possible for procedural learning to occur in the input and output phonology transductions; we might expect such tuning to be manifested in the form of repetition priming. This, of course, is precisely the pattern of behavioral results observed in patients with hippocampal amnesia. Such patients are virtually unable to learn new word meanings (e.g., Gabrieli, Cohen, & Corkin, 1988; Grossman, 1987) as a result of their impairment in declarative memory. However, these same patients exhibit intact repetition priming for both known and novel words (e.g., Haist, Musen, & Squire, 1991) as a result of their relatively spared procedural memory. The present view of procedural and declarative memory may therefore offer some insight into aspects of language learning: It suggests that the apparently unitary process of learning a new word is in fact based on two quite different forms of memory and learning.

Skill Learning and Repetition Priming: Constructs or Labels?

In our discussion throughout this article, we have been concerned with analyzing relationships between the standard measures of skill learning and repetition priming. We have attempted to show that observed dissociations between these measures do not constitute evidence that there are different underlying processes. Now we wish to offer a more radical view, which has three parts and a conclusion. The three parts of the view are as follows. (a) There are no canonical or "right" measures of either skill learning or repetition priming. (b) Underlying the effects that are measured as skill learning or repetition priming, there is only a single factor: the facilitatory effect of repetition of various aspects of stimuli on task performance. This facilitation has its effects at every level at which there is repetition. (c) Skill learning is just a form of repetition effect.

The conclusion is that "skill learning" and "repetition priming" are best viewed as terminological labels rather than as psychological constructs.

Let us begin by considering a digit-entering task in which participants are presented with a variety of different types of stimuli (five-digit strings) in each block. In particular, let us consider the types of stimuli that are summarized in Table 5 and assume that each block of the digit-entering task consists of presentation of an equal number of each of these types of stimulus.

Type A stimuli (see Table 5) are five-digit strings that do not repeat within or across blocks, and in which the two-digit transitions are random, both within and across blocks. These might be labeled "Random Unique" stimuli. Type B stimuli are five-digit strings that do not repeat within or across blocks, and which incorporate a fixed set of two-digit transitions within a block, which remains constant across blocks. These correspond to the "rule-following unique" stimuli of Poldrack et al. (1999, Experiment 1). Type C stimuli are five-digit strings that repeat across blocks, but which incorporate random two-digit transitions. These might be labeled "random repeating" stimuli. Type D stimuli are five-digit strings that repeat across blocks, and which incorporate a fixed set of two-digit transitions within a block, which remains constant across blocks; let us further assume that the particular two-digit transitions they incorporate are the same as those incorporated in the Type B stimuli. Type D stimuli also correspond to the "rule-following repeating" stimuli of Poldrack et al.

Let us now consider participants' (or the model's) performance on Type A stimuli (the random unique stimuli), over several blocks of performance in the digit-entering task. We would expect facilitation in performance on Type A stimuli, as depicted in Figure 18. Such facilitation would in an important sense be a repetition effect:

Table 5
Some Possible Types of Stimuli Presented in Each Block in a Digit-Entering Task

Type	Stimulus characteristics	Description
A	Nonrepeating five-digit transitions, randomly varying two-digit transitions	Random unique
B	Nonrepeating five-digit transitions, two-digit transitions that are systematic within a block, consistent across blocks	Rule-following unique
C	Repeating five-digit transitions, randomly varying two-digit transitions	Random repeating
D	Repeating five-digit transitions, two-digit transitions that are systematic within a block (and same as transitions in Type B stimuli), consistent across blocks	Rule-following repeating

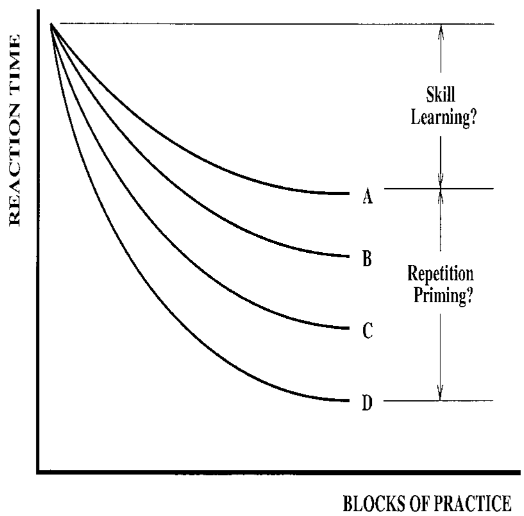


Figure 18. Pattern of performance that is expected on various types of stimuli in a digit-entering task.

Each time the system is exposed to a five-digit representation containing the digit *3* and in response executes a keystroke for the digit *3*, it becomes more practiced in making this mapping, and this is a repetition effect. Next consider performance on the Type B stimuli (the rule-following unique stimuli). Compared with the repetition of individual digits in Type A stimuli, the Type B stimuli incorporate repetition at a second level as well: There is systematic repetition of first-order transitions (i.e., two-digit transitions), both within blocks and across blocks. (Note that repetition of individual digits, as in Type A stimuli and all other stimuli, can be thought of as repetition of zero-order transitions between digits.) As a result of such repetition, the processing system would become practiced on the specific two-digit transitions in the Type B stimuli. We would therefore expect performance on Type B stimuli to be better than on Type A stimuli, as depicted in Figure 18. Next considering performance on Type C stimuli (random repeating stimuli), we would expect, as shown in Figure 18, that performance would be better than on Types A and B. This is because these stimuli incorporate repeating fourth-order digit transitions (i.e., entire five-digit transitions are repeated), and this greater degree of repetition should facilitate processing of Type C stimuli relative to the Type A and B stimuli. Finally, performance on Type D stimuli (rule-following repeating stimuli) should be best of all, because these stimuli incorporate repeating zero-order, first-order, and fourth-order transitions.

The question now arises of what is to be considered "skill learning." One proposal, illustrated in Figure 18, might be to use the term skill learning to refer to the facilitation in performance on Type A stimuli, because these have the smallest repetition component. However, as we have just argued, the facilitation in performance on these stimuli is no less a repetition effect than is facilitation in performance on any of the other types of stimuli. To see this point more clearly, let us imagine a transformation of the digit-entering task, called the artificial character-entering task. In this task, the stimuli are five-character strings composed of characters chosen from an artificial alphabet of 200 characters that are unfamilar to the participant. The computer system has been appropriately modified so that the screen can display these characters, and the keys on the keyboard are labeled with these characters. The participant's task is simply to type on the keyboard each five-character string that appears on the display, just as in the digit-entering task.

Now let us suppose that (unknown to the participant) we designate nine randomly chosen characters of this alphabet to be "digits." Four sets of stimuli that correspond to Types A–D above are constructed using only these nine characters. If we conducted the artificial character-entering task with these four kinds of stimuli in each block, we would expect the relative patterns of participants' performance on these types of stimuli to be very similar to the relative patterns of performance described above for the digit-entering task (although absolute levels of performance would undoubtedly be lower). Skill learning would be measured as improvement in performance on Type A stimuli, as proposed for the digit-entering task.

Let us now add a fifth kind of stimulus to the mix. These stimuli are made up of the "nondigit" characters. Furthermore, in these stimuli, no character ever repeats, that is, each of these "nondigit" characters appears exactly once throughout the experiment. Let us call these Type A′ stimuli. We would certainly expect performance to be worse on Type A′ stimuli than on Type A stimuli, so the curve for Type A′ stimuli would lie above that for Type A stimuli. Arguably, facilitation on Type A′ stimuli is now the better measure of skill learning than is facilitation on Type A stimuli, because the Type A′ stimuli incorporate even less repetition than the Type A stimuli. But this highlights the fact that there is repetition in the Type A stimuli and emphasizes that what is called "skill learning" is itself a repetition effect—in this case, the repetition of specific digits (real or artificial). The drawing of a sharp distinction between skill learning and repetition priming is therefore somewhat arbitrary. Moreover, there is not necessarily a canonical best measure of skill learning—it might always be possible to find some other stimulus type that contains even less repetition at some level of stimulus structure or processing in the task.

Similar considerations apply to the designation of any one particular measure as "repetition priming." Returning to the digit-entering task and Figure 18, one proposal might be to view the difference between performance on Type A and Type D stimuli as the appropriate measure of repetition priming because this difference reflects the maximum effect of repetition. This raises the question, however, of what to call the many other differences. To take just one example, the difference in performance between Type B and Type D stimuli is clearly due to the fact that there is greater repetition in the Type D than in the Type B stimuli. Is this difference to be viewed as something other than repetition priming? The same question applies to the difference between every pair of curves depicted in Figure 18. It might be argued that even if the difference between performance on Type A and Type D stimuli is not the only measure of repetition priming, it nevertheless remains in some sense the canonical or best measure of repetition priming. To examine this idea, let us consider an additional type of stimulus (Type D′), in which not only do the five-digit transitions and two-digit transitions repeat, but so do three-digit transitions. We can expect that performance on these stimuli would be better than on Type D stimuli. But if so, the canonical measure of repetition priming would now be the difference between performance on Type A and Type D′ stimuli. There appears to be no absolute measure of repetition priming that is “truly” the best.

We suggest that a more appropriate way to think about all of these effects is that they all reflect the effects of repetition at various levels of stimulus structure. The effect of repetition at the least repetitive of the levels that is measured in a given task is what might be labeled “skill learning.” However, it is really no different from any other repetition effect. Also, among the various levels of stimulus structure at which the effects of repetition are measured, there is no particular pair that is uniquely qualified to be thought of as indexing real “repetition priming.” The label can perhaps be applied to the difference between those two levels that exhibit the greatest and least degree of repetition. However, in reality, there is a multiplicity of repetition effects. Note also that these various repetition effects may each be distributed across a variety of processors, as discussed in the *Multiple Levels of Skill Learning and Repetition Priming* subsection. The effect of repetition at the level of two-digit stimulus structure, for example, may have a processing locus that is distributed across multiple transducers (i.e., levels of processing) and is not necessarily confined to one particular transducer. In sum, “skill learning” and “repetition priming” are merely labels that are applied to somewhat arbitrarily selected levels of repetition effect, as we have pointed out in previous work (Cohen & Eichenbaum, 1993; Cohen, Poldrack, & Eichenbaum, 1997).

Conclusions

As we have discussed in this article, the relationship between the phenomena of skill learning and repetition priming has been the subject of considerable debate, with some reports suggesting that they arise from a single mechanism (Cohen & Squire, 1980; Kirsner & Speelman, 1993; Logan, 1990) and other reports suggesting that they arise from different mechanisms (Kirsner & Speelman, 1996; Schwartz & Hashtroudi, 1991). We believe that this article makes a strong case that skill learning and repetition priming are manifestations of a single underlying procedural memory mechanism. We offered a particular characterization of such a procedural memory mechanism and showed how it can be related to the debate over memory systems, to the phenomena of skill acquisition and automaticity, and to aspects of language learning. At the very least, we believe that the present work helps clarify the terms of the debate over the relationship between skill learning and repetition priming. To the extent that our arguments are convincing, it may also represent a first step toward resolution of that debate and toward a broader conceptualization of procedural memory and its relationship to other domains of memory and cognition.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369–406.

Bartha, G. T., & Thompson, R. F. (1995). Cerebellum and conditioning. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 169–172). Cambridge, MA: MIT Press.

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 657–668.

Blaxton, T. A. (1995). A process-based view of memory. *Journal of the International Neuropsychology Society, 1,* 112–114.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18,* 355–387.

Bock, J. K., & Loebell, H. (1990). Framing sentences. *Cognition, 35,* 1–39.

Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99,* 150–171.

Buckner, R., Peterson, S. E., Ojemann, J. G., Miezin, F. M., Squire, L. R., & Raichle, M. E. (1995). Functional anatomical studies of explicit and implicit memory retrieval tasks. *Journal of Neuroscience, 15,* 12–29.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.

Chandler, J. P. (1969). Subroutine STEPIT: Finds local minima of a smooth function of several parameters. *Behavioral Science, 14,* 81–82.

Chang, F., Griffin, Z. M., Dell, G. S., & Bock, J. K. (1997, August). *Modeling structural priming as implicit learning.* Poster presented at the Conference on Computational Psycholinguistics, Berkeley, CA.

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 521–533.

Cleeremans, A. (1993). *Mechanisms of implicit learning: Connectionist models of sequence processing.* Cambridge, MA: MIT Press.

Cleeremans, A., Servan-Schreiber, D., & McClelland, J. (1989). Finite state automata and simple recurrent networks. *Neural Computation, 1,* 372–381.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing model of the Stroop effect. *Psychological Review, 97,* 332–361.

Cohen, N. J. (1984). Preserved learning capacity in amnesia: Evidence for multiple memory systems. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (pp. 83–103). New York: Guilford Press.

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system.* Cambridge, MA: MIT Press.

Cohen, N. J., Poldrack, R. A., & Eichenbaum, H. (1997). Memory for items and memory for relations in the procedural/declarative memory framework. *Memory, 5,* 131–178.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science, 210,* 207–209.

Cousineau, D., & Larochelle, S. (1997). PASTIS: A program for curve and distribution analysis. *Behavioral Research Methods, Instruments, and Computers, 29,* 542–548.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science, 23,* 517–542.

Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science, 17,* 149–195.

Demb, J. B., Desmond, J. E., Wagner, A. D., Stone, M., Lee, A. T., Glover, G. H., & Gabrieli, G. D. E. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *Journal of Neuroscience, 15,* 5870–5878.

Dollaghan, C. (1985). Child meets word: “Fast mapping” in preschool children. *Journal of Speech and Hearing Research, 28,* 449–454.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14,* 179–211.

Fendrich, D. W., Healy, A. F., & Bourne, L. E. (1991). Long-term repetition effects for motoric and perceptual procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 137–151.

Fleischman, D. A., Gabrieli, J. D. E., Reminger, S., Rinaldi, J., Morrell, F., & Wilson, R. (1995). Conceptual priming in perceptual identification for patients with Alzheimer's disease and a patient with right occipital lobectomy. *Neuropsychology, 9,* 187–197.

Gabrieli, J. D. E. (1995). A systematic view of human memory processes. *Journal of the International Neuropsychology Society, 1,* 115–118.

Gabrieli, J. D. E., Cohen, N. J., & Corkin, S. (1988). The impaired learning of semantic knowledge following bilateral medial temporal-lobe resection. *Brain, 7,* 157–177.

Gabrieli, J. D. E., Fleischman, D. A., Keane, M. M., Reminger, S. L., & Morrell, F. (1995). Double dissociation between memory systems underlying explicit and implicit memory in the human brain. *Psychological Science, 6,* 76–82.

Graf, P., & Schacter, D. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 501–518.

Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of Cognitive Neuroscience, 7,* 497–510.

Grafton, S. T., Mazziota, J. C., Presty, S., Friston, K., Frackowiak, R. S. J., & Phelps, M. E. (1992). Functional anatomy of human procedural learning determined with regional cerebral blood flow and PET. *Journal of Neuroscience, 12,* 2542–2548.

Grossman, M. (1987). Lexical acquisition in alcoholic Korsakoff psychosis. *Cortex, 23,* 631–644.

Gupta, P., & Dell, G. S. (1999). The emergence of language from serial order and procedural memory. In B. MacWhinney (Ed.), *28th Carnegie Mellon Symposium on Cognition: The emergence of language.* Hillsdale, NJ: Erlbaum.

Gupta, P., & MacWhinney, B. (1992). Integrating category acquisition with inflectional marking: A model of the German nominal system. *Proceedings of the 14th Annual Conference of the Cognitive Science Society* (pp. 253–258). Hillsdale, NJ: Erlbaum.

Gupta, P., & Touretzky, D. S. (1994). Connectionist models and linguistic theory: Investigations of stress systems in language. *Cognitive Science, 18,* 1–50.

Haist, F., Musen, G., & Squire, L. R. (1991). Intact priming of words and nonwords in amnesia. *Psychobiology, 19,* 275–285.

Hazeltine, E., Grafton, S. T., & Ivry, R. (1997). Attention and stimulus characteristic determine the locus of motor-sequence encoding. *Brain, 120,* 123–140.

Heindel, W. C., Salmon, D. P., Shults, C. W., Wallicke, P. A., & Butters, N. (1989). Neuropsychological evidence for multiple implicit memory systems: A comparison of Alzheimer's, Huntington's, and Parkinson's disease patients. *Journal of Neuroscience, 9,* 582–587.

Hoeffner, J. (1992). Are rules a thing of the past? Learning the past tense in an attractor network. *Proceedings of the 14th Annual Conference of the Cognitive Science Society* (pp. 861–866). Hillsdale, NJ: Erlbaum.

Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior, 22,* 485–508.

Jacoby, L. L., & Kelley, C. M. (1992). A process-dissociation framework for investigating unconscious influences: Freudian slips, projective tests, subliminal perception and signal detection theory. *Current Directions in Psychological Science, 1,* 175–179.

Jordan, M. I. (1986). *Serial order: A parallel distributed processing approach* (Rep. No. 8604). La Jolla: University of California, San Diego, Institute for Cognitive Science.

Jordan, M. I., & Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science, 16,* 307–354.

Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature, 377,* 155–158.

Kirsner, K. (1998). Implicit memory. In K. Kirsner, C. Speelman, M. Mayberry, A. O'Brien-Malone, M. Anderson, & C. Macleod (Eds.), *Implicit and explicit mental processes.* Hillsdale, NJ: Erlbaum.

Kirsner, K., & Speelman, C. (1993). Is lexical processing just an ACT? In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory.* Hillsdale, NJ: Erlbaum.

Kirsner, K., & Speelman, C. (1996). Skill acquisition and repetition priming: One principle, many processes? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 563–575.

Kolers, P. A. (1975). Specificity of operations in sentence recognition. *Cognitive Psychology, 7,* 289–306.

Levelt, W. J. M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology, 14,* 78–106.

Logan, G. D. (1978). Attention in character classification: Evidence for the automaticity of component stages. *Journal of Experimental Psychology: General, 107,* 32–63.

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 189–207.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology, 22,* 1–35.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition, 40,* 121–157.

McAndrews, M. P., & Moscovitch, M. (1990). Transfer effects in implicit tests of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 772–788.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102,* 419–457.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114,* 159–188.

McClelland, J. L., & Rumelhart, D. E. (Eds.). (1986). *Parallel distributed processing: Vol. 2. Psychological and biological models.* Cambridge, MA: MIT Press.

McClelland, J. L., & Rumelhart, D. E. (1988). *Explorations in parallel distributed processing.* Cambridge, MA: MIT Press.

McClelland, J. L., & St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes, 4,* 287–335.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 24, pp. 109–165). New York: Academic Press.

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J. McGaugh, & N. Weinberger (Eds.), *Neurobiology of learning and memory* (pp. 65–77). New York: Guilford Press.

Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience, 4,* 257–267.

Moscovitch, M. (1994). Memory and working with memory: Evaluation of a component process model and comparisons with other models. In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 269–310). Cambridge, MA: MIT Press.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition.* Hillsdale, NJ: Erlbaum.

Nissen, M. J., Knopman, D. S., & Schacter, D. L. (1987). Neurochemical dissociation of memory systems. *Neurology, 37,* 789–794.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Poldrack, R. A. (1995). *The relationship between skill learning and repetition priming.* Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.

Poldrack, R. A., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). The neural basis of visual skill learning: An fMRI study of mirror reading. *Cerebral Cortex, 8,* 1–10.

Poldrack, R. A., Selco, S. A., Field, J. E., & Cohen, N. J. (1999). The relationship between skill learning and repetition priming: Experimental and computational analyses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 208–235.

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

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variation in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning: II. Theory and research.* New York: Appleton-Century-Crofts.

Roediger, H. L., & Blaxton, T. A. (1987). Retrieval modes produce dissociations in memory for surface information. In D. S. Gorfein & R. R. Hoffman (Eds.), *Memory and cognitive processes: The Ebbinghaus Centennial Conference.* Hillsdale, NJ: Erlbaum.

Roediger, H. L. I., Buckner, R., & McDermott, K. B. (1999). Components of processing. In J. K. Foster & M. Jelicic (Eds.), *Memory: Systems, process, or function?* (pp. 31–65). New York: Oxford University Press.

Roediger, H. L. I., & McDermott, K. B. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63–131). Amsterdam: Elsevier.

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In D. Rumelhart & J. McClelland (Eds.), *Parallel distributed processing: Vol. 1. Foundations* (pp. 318–362). Cambridge, MA: MIT Press.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Vol. 1. Foundations.* Cambridge, MA: MIT Press.

Rumelhart, D. E., & McClelland, J. L. (1987). Learning the past tenses of English verbs: Implicit rules or parallel distributed processes? In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 195–248). Hillsdale, NJ: Erlbaum.

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

Schacter, D., Chiu, C. Y. P., & Ochsner, K. N. (1993). Implicit memory: A selective review. *Annual Review of Neuroscience, 16,* 159–182.

Schacter, D. L., & Buckner, R. L. (1998). On the relations among priming, conscious recollection, and intentional retrieval: Evidence from neuroimaging research. *Neurobiology of Learning and Memory, 70,* 284–303.

Schacter, D. L., & Church, B. A. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 915–930.

Schacter, D. L., Wagner, A. D., & Buckner, R. L. (2000). Memory systems of 1999. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 627–643). New York: Oxford University Press.

Schmajuk, N. A. (1995). Conditioning. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 238–243). Cambridge, MA: MIT Press.

Schneider, W., & Fisk, A. D. (1982). Degree of consistent training: Improvements in search performance and automatic process development. *Perception and Psychophysics, 31,* 160–168.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66.

Schwartz, B. L., & Hashtroudi, S. (1991). Priming is independent of skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 1177–1187.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20,* 11–21.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting memory and learning. *Journal of Cognitive Neuroscience, 4,* 232–243.

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology, 44,* 453–495.

Squire, L. R., Ojemann, J. G., Miezin, F. M., Petersen, S. E., Videen, T. O., & Raichle, M. E. (1992). Activation of the hippocampus in normal humans: A functional anatomical study of memory. *Proceedings of the National Academy of Sciences, USA, 89,* 1837–1841.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review, 88,* 135–170.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory.* New York: Academic Press.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247,* 301–306.

Vaidya, C. J., Gabrieli, J. D. E., Verfaellie, M., Fleischman, D., & Askari, N. (1998). Font-specific priming following global amnesia and occipital lobe damage. *Neuropsychology, 12,* 183–192.

Wagner, A. D., Desmond, J. E., Demb, J. B., Glover, G. H., & Gabrieli, J. D. E. (1997). Semantic repetition priming for verbal and pictorial knowledge: A functional MRI study of left inferior prefrontal cortex. *Journal of Cognitive Neuroscience, 9,* 714–726.

Widrow, G., & Hoff, M. (1960). Adaptive switching circuits. *Institute of Radio Engineers, Western Electric Show and Convention, Convention Record, Pt. 4* (pp. 96–104).

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology, 8,* 227–233.

Witherspoon, D., & Moscovitch, M. (1989). Stochastic independence between two implicit memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 22–30.

(*Appendix follows*)

Appendix

Further Analysis of Previous Results

Schwartz and Hashtroudi (1991), Experiment 2

In their second experiment, Schwartz and Hashtroudi (1991, Experiment 2) examined whether the amount of skill learning in the partial-word identification task would influence the magnitude of priming in that task. Some of their participants were given five trials of preexperimental practice with identification of degraded words (note that Schwartz and Hashtroudi use the term "trials" to mean blocks), whereas other participants were given no preexperimental practice in partial-word identification. In the experiment itself, all participants first underwent a study phase in which they were given a new list of words to study, in undegraded form. All participants then underwent a test phase, in which they were required to perform partial-word identification, with the stimuli consisting of the studied words as well as new, nonstudied words, all stimuli being in degraded form. Schwartz and Hashtroudi reasoned that if priming effects are related to the acquisition of skills and procedures, then priming should be greater for those participants who had previous exposure to degraded words (i.e., skilled participants) than for those who did not (i.e., unskilled participants). Schwartz and Hashtroudi found that the magnitude of priming did not differ significantly for the two groups of participants and concluded that priming in partial-word identification is unrelated to the amount of previous practice in identifying degraded words.

Note that Schwartz and Hashtroudi (1991) used two different paradigms to study priming in the partial-word identification task. In Experiment 1, they used the multiple-repetition paradigm that we have examined extensively in this article, in which both repeating and unique items are presented in each of several blocks ("trials," in Schwartz and Hashtroudi's terminology). In Experiment 2, they used a study–test paradigm, which has two phases: a study phase, in which stimuli are presented outside the context of the task in which priming is to be studied; and a test phase, in which the studied items are presented once, along with nonstudied items, in the experimental task, in which priming is sought to be examined. Each of these paradigms has been used extensively in priming studies (examples of use of the multiple-repetition paradigm include Kirsner & Speelman, 1996; Poldrack et al., 1999, and Schwartz & Hashtroudi, 1991, Experiment 1; examples of use of the study–test paradigm include Church & Schacter, 1994, and Schacter & Church, 1992). It is of some interest to examine the relationship between these two paradigms, and we will begin our reanalysis of Schwartz and Hashtroudi's Experiment 2 results by attempting to relate the paradigms. Figure A1 redisplays the performance on unique stimuli and performance on repeating stimuli curves from the partial-word identification task in Schwartz and Hashtroudi's Experiment 1. In addition, the figure displays the results of Schwartz and Hashtroudi's Experiment 2. Let us consider how the two paradigms relate to each other.

The performance of the nonpracticed participants in Experiment 2 (the study–test paradigm) can be related to Trial 1 performance in Experiment 1 (the multiple-repetition paradigm). In the test phase of Experiment 2, nonpracticed participants performed the partial-word identification task for the first time, as did participants in Experiment 1 on Trial 1; these are corresponding data points in the two tasks. This is corroborated by comparison of skill levels: For the Experiment 2 nonpracticed participants, performance on nonstudied stimuli in the test phase is very similar to the performance of Experiment 1 participants on novel stimuli in Trial 1. For this reason, we have plotted the performance of Experiment 2 nonpracticed participants at the point on the x-axis that corresponds to Trial 1 in Experiment 1. The performance of the practiced participants in the test phase in Experiment 2 can also be mapped onto the structure of Experiment 1. The Experiment 2 practiced participants had received five trials of practice (on nonstudied stimuli) in the partial-word identification task prior to the study and test phases. At the test phase, therefore, these participants were at a level of skill that was equivalent to Experiment 1 participants at Trial 6; these are corresponding data points in the two tasks. This is corroborated by comparison of skill levels: For the Experiment 2 practiced participants, performance on nonstudied items is very similar to the performance of Experiment 1 participants on unique stimuli at Trial 6. We have therefore graphed the results of the Experiment 2 practiced participants at the point on the x-axis that corresponds to Trial 6 in Experiment 1. These observations apply not only to the present experiments, but also establish a correspondence between the structure of multiple-repetition paradigms and study–test paradigms in general.

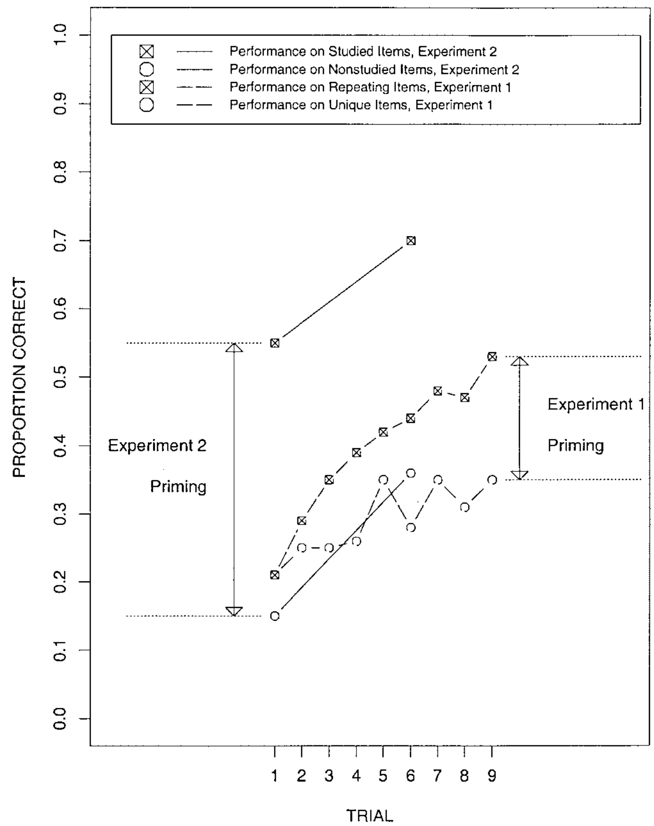


Figure A1. Comparison of partial-word identification task in two different priming paradigms (Schwartz & Hashtroudi, 1991, Experiments 1 and 2).

It is also worth examining the magnitude of the priming effect, which is much greater in Experiment 2 than it is in Experiment 1. In general, it is to be expected that priming in a study–test paradigm will be greater than it is in Trial 1 of a multiple-repetition paradigm.[A1] However, this does not explain why priming in Experiment 2 is so much greater than even the

[A1] To see this, note that in a multiple-repetition paradigm, there is no difference from the participant's point of view between unique and repeating stimuli: In Trial 1, they are all unique stimuli. In the test phase of a study–test paradigm, however, there is by definition a difference between studied and nonstudied items, even if participants have had no practice in the task that they perform in the test phase.

eventual magnitude of priming in Experiment 1 at the end of nine trials. The reason for this appears to be that the study phase in Experiment 2 exposed participants to the studied items "in the clear," whereas in Experiment 1, participants were only exposed to repeating items in their degraded form.

This bring us, finally, to the crucial question: In Experiment 2, why is priming for the practiced participants no greater than is priming for the nonpracticed participants? If skill learning is related to priming, then shouldn't the more skilled participants exhibit more priming? The answer to this question is in the negative. Let us compare performance on Trial 1 and on Trial 6 in the multiple-repetition paradigm. There are two differences between performance at these two levels. First, participants have had more practice on unique stimuli at Trial 6 than at Trial 1. Because of this, performance on unique stimuli is superior at Trial 6 than at Trial 1; that is, there has been an increase in skill. Second, participants have had more practice on the specific repeating stimuli at Trial 6 than at Trial 1. Because of this, performance on these specific repeating stimuli has improved more than generalized performance has improved; consequently, repetition priming has increased. Next, let us consider the difference between the nonpracticed and the practiced participants in the test phase of the study–test paradigm. Practiced participants have had five trials of practice on nonstudied (i.e., unique) stimuli. They are at a Trial 6 level of performance, with respect to general ability, that is, skill. This skill level is superior to that of nonpracticed participants, which is in effect at a Trial 1 level. However, practiced participants have had no more exposure to the studied stimuli than nonpracticed participants have had. Therefore, there is no reason why the advantage for studied over nonstudied stimuli should be any different for the two groups. The crucial point is that practiced participants in the study–test paradigm have had five more trials of practice (than have the nonpracticed participants) on the nonstudied stimuli, but not on the studied stimuli. With respect to studied items, there is no difference between practiced and nonpracticed participants in the study–test paradigm, unlike the situation in the multiple-repetition paradigm, where there is a difference between participants at Trial 6 and at Trial 1, even with respect to the studied (i.e., repeating) items.

The expectation that practiced participants should show greater priming arises from two errors in reasoning. First of all, it fails to take into account the differences between the study–test and multiple-repetition paradigms. In a multiple-repetition paradigm, participants receive practice as well as further exposure to the repeating stimuli on each trial. Thus participants' practice on the task increases along with the number of times they have been exposed to the repeating stimuli; extent of practice and number of repetitions are tightly coupled. In such a paradigm, it will therefore often be the case that both skill learning and repetition priming increase together (although this is not a necessary result, as we showed in the *Patterns of Increase and Decrease* subsection). In the study–test paradigm, however, practice and repetition do not increase together. In Experiment 2, the practice that was given to participants was not accompanied by practice on to-be-studied items. The effects of practice and of repetition were therefore decoupled. The expectation that practiced participants in Experiment 2 should have shown greater priming than should nonpracticed participants thus arises from an erroneous transfer of expectations from a multiple-repetition paradigm to the study–test paradigm. There is also, however, a second error in reasoning: the expectation that skill learning and repetition priming should necessarily increase together in a multiple-repetition paradigm is itself mistaken (as we showed in the *Patterns of Increase and Decrease* subsection). All possible patterns of increase and decrease in skill learning and repetition priming are consistent with a single underlying mechanism, and there is no reason to expect that any particular one of these patterns must hold. For all of these reasons, the results of Schwartz and Hashtroudi's Experiment 2 do not constitute evidence that skill learning and repetition priming are supported by different mechanisms.

Schwartz and Hashtroudi (1991), Experiment 3

Schwartz and Hashtroudi (1991) also reported a further study aimed at examining whether skill learning and repetition priming are supported by the same or different mechanisms. In Experiment 3, they examined whether word frequency would modulate priming. The rationale was that participants have greater preexperimentally acquired skill in processing high-frequency words compared with low-frequency words. If skill learning and repetition priming are related, then priming should be greater for the high-frequency words than for the low-frequency words. Experiment 3 used the partial-word identification task in a multiple-repetition paradigm. Word frequency was manipulated between participants. Schwartz and Hashtroudi found that skill learning was significantly greater for the high-frequency words than for the low-frequency words, whereas the magnitude of repetition priming was unaffected by frequency. They suggested that these differential effects of word frequency indicate that skill learning and repetition priming are unrelated (p. 1183).

The data from Experiment 3 are redrawn in Figure A2a and A2b for low-frequency and high-frequency words, respectively. The results indicate that performance on uniques and performance on repeats is similarly affected by word frequency. Thus performance on unique stimuli is higher for high-frequency words than for low-frequency words, and similarly, performance on repeating stimuli is higher for high-frequency words than for low-frequency words. Additionally, performance on unique stimuli increases more with practice for high-frequency words than for low-frequency words (i.e., there is more skill learning for high-frequency words than for low-frequency words); and similarly, performance on repeating stimuli increases more with practice for high-frequency words than for low-frequency words (i.e., there is a greater benefit of repetition for high-frequency words than for low-frequency words).

As can also be seen in Figure A2, the gap between the performance on unique stimuli and performance on repeating stimuli curves is very similar for both high- and low-frequency words; that is, repetition priming is similar. This can also be seen from the fact that the repetition priming functions do not appear to differ systematically for the high-frequency and low-frequency words. As a result, the derived repetition priming and skill learning measures show a dissociation: Skill learning is greater for high-frequency words than for low-frequency words, whereas repetition priming is the same for high- and low-frequency words. In light of our analyses, we can see that this is an artifactual result. It follows from the definitions of the skill learning and repetition priming measures. The results of Schwartz and Hashtroudi's Experiment 3 therefore do not constitute evidence for the unrelatedness of processing underlying skill learning and repetition priming. Here, once again, as in other cases we have discussed, a clearer understanding of the pattern of results can be achieved by taking into account the performance on unique stimuli and performance on repeating stimuli curves. When these measures are incorporated into the analysis, it becomes clear that word frequency has almost identical effects on the processing of unique and repeating stimuli in the partial-word identification task. What the results of Experiment 3 suggest, in fact, is that the processing of unique and repeating stimuli is highly related.

McAndrews and Moscovitch (1990), Experiment 4

Let us also consider results reported by McAndrews and Moscovitch (1990, Experiment 4), who examined participants' anagram-solving ability. In the "study" or "training" phase in that experiment, participants were given 60 five-letter anagrams to solve. All of these anagrams incorporated a specific reordering pattern such that if the original word was composed of letters in the order "12345", the corresponding anagram consisted of letters in the order "51324". Over the course of training on these 60 anagrams, participants showed steady improvement (with solution time following a decreasing power function), which suggested that they were learning the skill of solving anagrams that followed this ordering pattern. One week after the training session, the same participants underwent the

(*Appendix continues*)

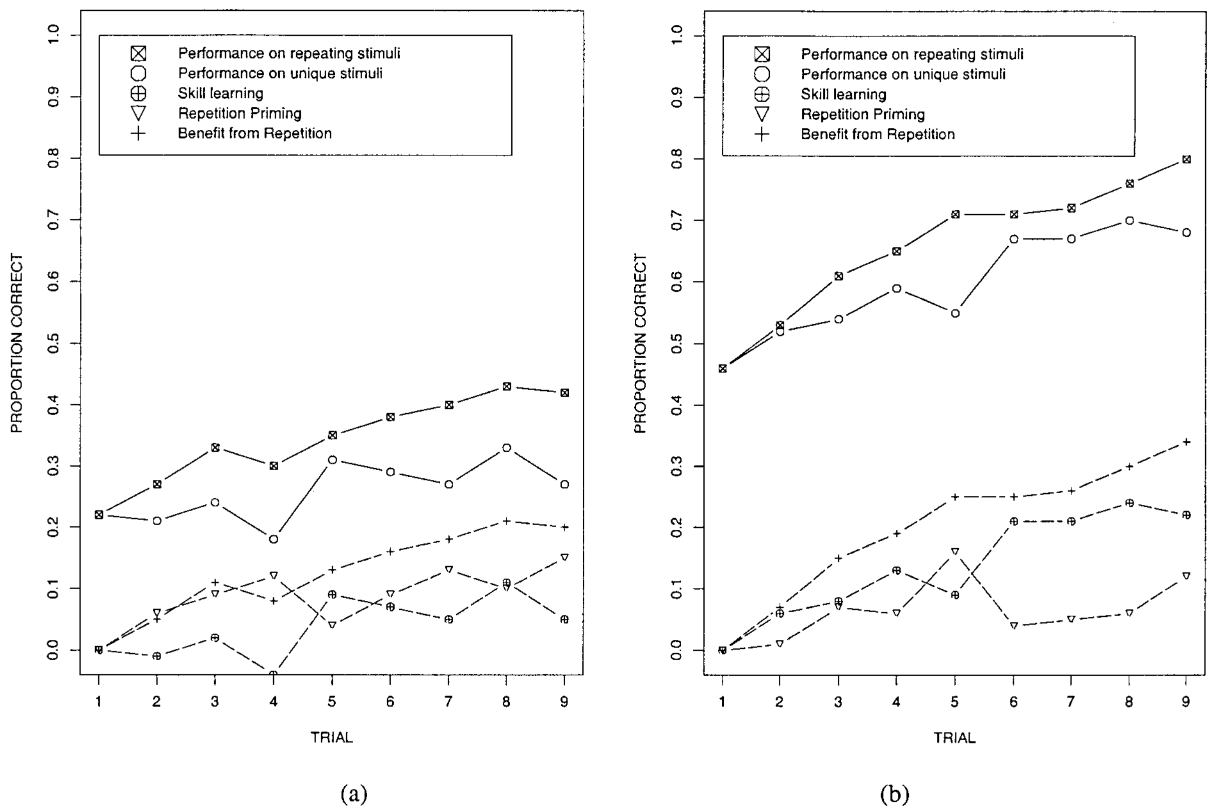


Figure A2. Results from Schwartz and Hashtroudi (1991, Experiment 3). (a) Low-frequency words. (b) High-frequency words.

"test" phase of the experiment, in which they were presented with a further 100 five-letter anagrams. These 100 anagrams consisted of four sets of 25 anagrams each. In one set (which we will term the rule-following repeats), the anagrams were identical to 25 of the anagrams presented during training. In a second set (which we will term the non-rule-following repeats), the stimuli were anagrams of 25 of the same words presented at study, but were derived using a different reordering scheme. A third set of 25 anagrams (which we will term the rule-following uniques) were novel and incorporated the same "51324" reordering rule as in the training stimuli. The fourth set of 25 anagrams (which we will term the non-rule-following uniques) consisted of novel anagrams that incorporated a different reordering scheme from that used in the training stimuli.

The test phase thus used a 2 × 2 factorial design, with the two factors being Presentation (whether or not the anagram was a repetition, as compared with the training phase) and Rule Status (whether or not the anagram followed the same reordering rule as in the training phase). McAndrews and Moscovitch found that although there was a main effect of both Presentation and Rule Status, there was no significant interaction. Thus, both the application of a previously learned implicit rule (i.e., the learning of a skill) and previous experience with a specific anagram (i.e., the effect of repetition) facilitated performance, and the effects of these two sources of facilitation were additive and independent. The authors took this lack of interaction to be inconsistent with the view that skill learning and repetition priming are supported by the same procedural memory system (p. 784).

Our analytical framework makes it possible to see that such statistical independence is not very informative as to the underlying processing. Let us suppose that participants were trained on each of these kinds of stimuli for several blocks of practice. Figure A3a illustrates four performance functions of the kind that we might expect for the four categories of stimuli that were presented in the test phase of McAndrews and Moscovitch's Experiment 4. With practice, we would expect response times to (perhaps) decrease somewhat for non-rule-following unique anagrams, to decrease somewhat more for rule-following unique anagrams, decrease still more for non-rule-following repeat anagrams, and to decrease the most for the rule-following repeat anagrams. The general shape assumed for these functions is consistent with the results reported by McAndrews and Moscovitch from the experiment's training phase (in which performance on rule-following repeat anagrams followed the power law of practice).

From our discussion throughout this article, and particularly from our presentation of simulation results and from our discussion in the *Skill Learning and Repetition Priming: Constructs or Labels*? subsection, it should be clear that a family of four performance functions of this kind is quite consistent with the operation of a single underlying learning mechanism. Now, the data points reported by McAndrews and Moscovitch can be seen as representing four points, one on each of these functions. In Figure A3a, we have labeled these points NRU (non-rule unique), RU (rule unique), NRR (non-rule repeat),

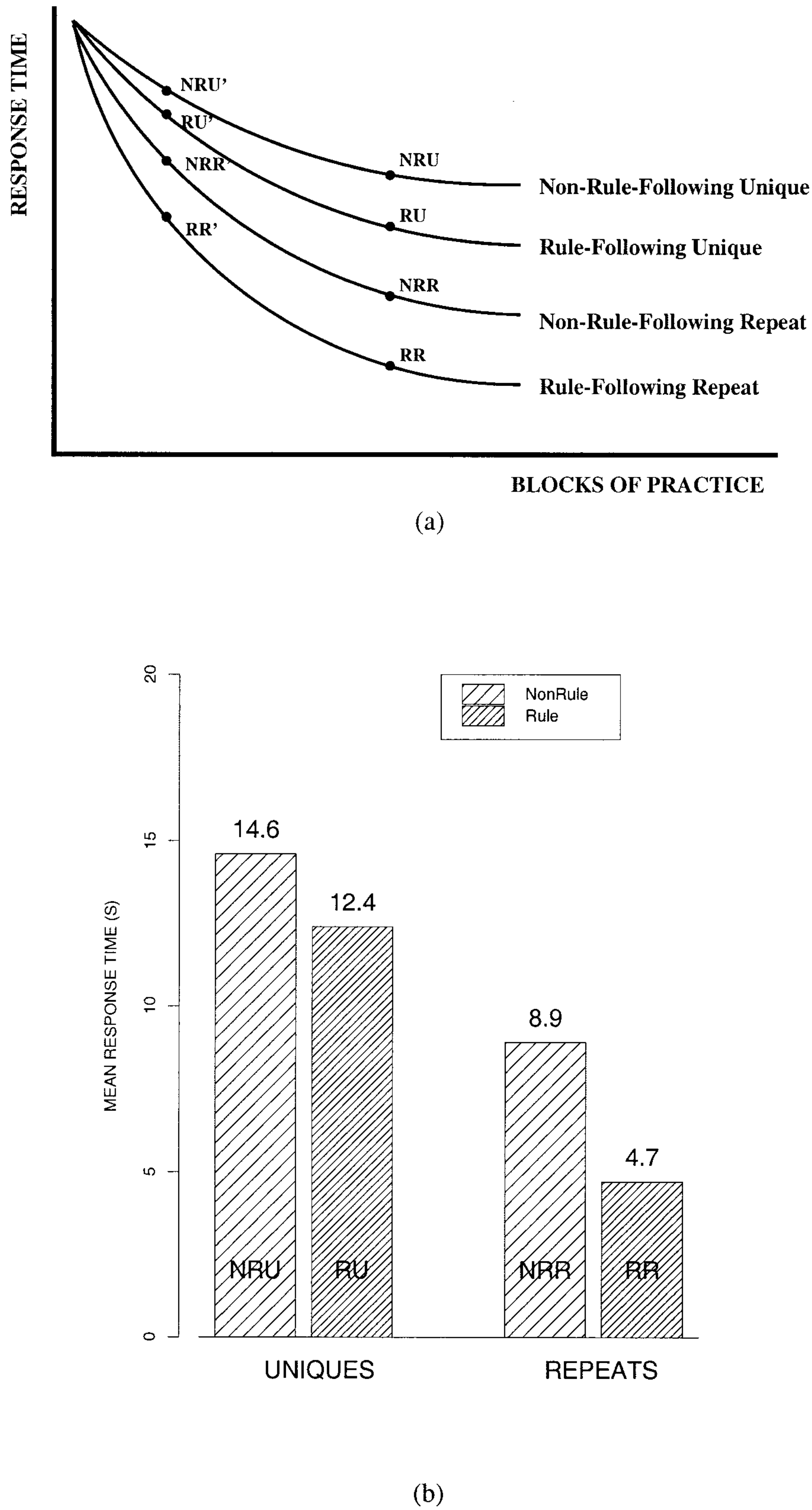


Figure A3. Analysis of results from McAndrews and Moscovitch (1990, Experiment 4). (a) Performance functions expected for various kinds of stimuli used at test. (b) Results for the four data points for unaware participants (redrawn from McAndrews & Moscovitch, 1990, Table 9). NRU = non-rule-following unique; RU = rule-following unique; NRR = non-rule-following repeat; RR = rule-following repeat.

(*Appendix continues*)

and RR (rule repeat). In Figure A3b, we plot the mean values reported by McAndrews and Moscovitch for each of these points. (These are data for participants who were unaware of the implicit anagram-ordering rule.)

McAndrews and Moscovitch's key finding was that there was no interaction between the rule and repetition effects, and they interpreted this as indicating a lack of commonality in underlying mechanisms. However, when we situate the four data points against the framework of our analysis, as in Figure A3b, it becomes clear that what we are trying to interpret is the relative magnitude of the differences between two pairs of performance functions at a particular point in practice. The lack of interaction in Figure A3b is in effect a finding that the separation between the rule-following unique and non-rule-following unique functions was not statistically different from the separation between the rule-following repeating and non-rule-following repeating functions at the level of practice at which participants were examined. Although there might not have been an interaction at that particular level of practice, there almost certainly would be an interaction at some other level of practice, given that the performance functions are power functions, as indicated by McAndrews and Moscovitch's own data. This can be seen by considering the points labeled NRU′, RU′, NRR′, and RR′ on the same performance curves. Thus, success or failure to find statistical differences at arbitrarily selected levels of practice says little about the underlying mechanisms that generate the functions. Here again, our theoretical framework helps clarify that these empirical results do not directly bear on the question of underlying mechanism.

Received August 15, 1997
Revision received April 30, 2001
Accepted May 22, 2001 ■