Journal of Experimental Psychology:
Learning, Memory, and Cognition
2005, Vol. 31, No. 1, 9–23

0278-7393/05/$12.00 DOI: 10.1037/0278-7393.31.1.9

The Influence of Grammatical, Local, and Organizational Redundancy on Implicit Learning: An Analysis Using Information Theory

Randall K. Jamieson and D. J. K. Mewhort
Queen's University at Kingston

People behave as if they know the structure of their environment. Because people rarely study that structure explicitly, several theorists have postulated an implicit learning system that abstracts that structure automatically. An alternative view is that people respond to local structure that derives from global structure. Measures are developed that quantify structure in a set of stimuli, in individual stimuli, and in encoded stimuli. The authors apply the measures to examine serial recall for sequences of colors generated using a stationary Markov grammar. They demonstrate that the 3 kinds of redundancy are confounded and show that the memorial advantage for grammatical stimuli reflects participants' use of local expressions of grammatical structure to aid learning.

People often behave as though they know the structure of their environment. Because people rarely study that structure explicitly, several theorists have taken sensitivity to structure as evidence for a specialized learning system, one that abstracts structure automatically and that guides behavior adaptively (Dienes, Broadbent, & Berry, 1991; Knowlton & Squire, 1992, 1994, 1996; Manza & Reber, 1997; Mathews et al., 1989; Reber, 1967, 1969, 1989, 1993).

In Reber's (1967) now classic study of implicit learning, participants recalled strings of letters that were constructed at random or according to rules (i.e., a Markov grammar). Early in the experiment, participants in both conditions took the same mean number of trials to recall both the grammatical and the random strings. After the first eight strings, however, participants who learned grammatical strings continued to improve, whereas those who learned random strings did not. Reber argued that the advantage for grammatical stimuli reflected participants' knowledge of the grammar. Because they were unable to articulate the rules of the grammar, he argued that the knowledge was implicit.

Critics of the implicit learning position have argued that the classic data are insufficient to force the claim that subjects abstracted the grammar. Brooks and his colleagues (Brooks, 1978; Brooks & Vokey, 1991; Vokey & Brooks, 1992, 1994; Whittlesea & Dorken, 1993; Whittlesea & Wright, 1997; Wright & Whittlesea, 1998) suggested, for example, that the similarity between a current stimulus and studied stimuli, whether physical or analogical, partially explains the advantage for grammatical stimuli: The grammatical constraints force similarities among the stimuli, and, because already learned instances are similar to current instances, knowledge from prior learning can be applied to current learning. Brooks and colleagues' explanation does not require that subjects know the grammar and is consistent with participants' inability to articulate the rules of the grammar.

Perruchet and his colleagues (e.g., Johnstone & Shanks, 2001; Perruchet & Gallego, 1997; Perruchet & Pacteau, 1990) offered a related critique and argued that participants organize the stimuli into higher order subjective units and then recode the original stimulus as a string of those subjective units. Servan-Schreiber and Anderson (1990) have developed a computational version of the theory.

Implicit learning has also been examined through classification tasks. In these experiments, participants study strings of symbols that are constructed with a grammar and then attempt to discriminate novel grammatical from novel ungrammatical cases (i.e., cases that are consistent with the grammar from those that are not). Although participants can discriminate grammatical strings, they cannot articulate the grammar. The result has been used as evidence in three main lines of argument:

1. Abstractionist theories propose that participants internalize the grammar at study and then use that knowledge to discriminate strings at test: The internalized grammar need be only a partial set of the grammatical rules. Strings that match the internalized version of the grammar are endorsed as grammatical, and those that do not are rejected. Because participants are unable to articulate the grammar, knowledge of the grammar is thought to be implicit (Knowlton & Squire, 1994, 1996; Mathews et al., 1989; McAndrews & Moscovitch, 1985; Reber, 1989, 1993).

2. Statistical theories propose that participants learn the regularities in studied items and then endorse strings that exhibit those regularities. For example, subjects might learn the bigrams in studied items and then endorse test

Randall K. Jamieson and D. J. K. Mewhort, Department of Psychology, Queen's University, Kingston, Ontario, Canada.

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada to D. J. K. Mewhort. The experiments reported in this article were presented at the annual meeting of the Canadian Society of Brain, Behaviour, and Cognitive Science, June 2004, St. John's, Newfoundland, Canada.

We are indebted to Beth Johns and Mike Jones for helpful comments on the article.

Correspondence concerning this article should be addressed to Randall K. Jamieson, Department of Psychology, Queen's University, Kingston, Ontario K7L 3N6, Canada. E-mail: randy@minerva.psyc.queensu.ca

strings that exhibit those bigrams (Perruchet & Gallego, 1997; Perruchet & Pacteau, 1990). Discrimination is also influenced by the positional constraints of bigrams in the studied items (i.e., whether remembered bigrams occurred at the beginning, middle, or end of studied strings). The statistical theories split on whether knowledge of the statistical regularities is implicit (Dienes et al., 1991; Dulany, Carlson, & Dewey, 1984; Gomez & Schvaneveldt, 1994; Knowlton & Squire, 1992, 1994, 1996; Mathews et al., 1989; Meulemans & Van der Linden, 2003; Servan-Schreiber & Anderson, 1990).

3. Processing theories propose that learning is strongly affected by how participants process the stimuli at both study and test. At study, participants encode a version of each studied stimulus that reflects the organization imposed on it. For example, a number string *7354* could be remembered as *7 3 5 4*; *73* and *54*; *7,354*; and so on. At test, stimuli that match the encoded versions of the studied stimuli are endorsed as grammatical. Because memory for studied stimuli is shaped by the encoding operations performed on them, discrimination reflects a processing-specific pattern of performance. The principles of encoding specificity (Tulving & Thomson, 1973) and transfer appropriate processing (Morris, Bransford, & Franks, 1977) are central to the perspective (Brooks, 1978; Brooks & Vokey, 1991; Higham, 1997; Johnstone & Shanks, 2001; Vokey & Brooks, 1992, 1994; Whittlesea & Dorken, 1993; Whittlesea & Wright, 1997; Wright & Whittlesea, 1998).

The abstractionist and the statistical theories both argue that performance reflects participants' learning of the regularities in a set of studied stimuli. The proposition is attractive because experimenters can ignore complexities based on variations in processing (e.g., encoding variability) and develop objective and parsimonious theories based on the properties of experimental stimuli. However, the processing theorists have argued that the abstractionist and statistical theories are incomplete, because they cannot explain changes in discrimination based on processing distinctions.

To make the point, Higham (1997) had participants study letter strings constructed with an artificial grammar. Afterward, the participants were told that the strings were constructed according to rules and then attempted to discriminate novel grammatical from novel ungrammatical strings.

Discrimination was tested when the study and test strings were instantiated through the use of (a) only consonants and (b) both consonants and vowels. If participants abstract the statistical regularities in the stimuli, performance should not differ in the two conditions—both stimulus sets were identical in bigram frequency. Higham (1997) proposed, however, that the materials would be remembered as letter strings in the first case and as sounds in the latter. Because the strings and sounds are coded in different ways, memory for the stimuli differs and, therefore, should change how strings are classified.

The prediction was confirmed, and Higham (1997) argued that the example shows that abstractionist and statistical theories provide incomplete explanations of structural learning because they fail to capture the processing-specific control on performance. Similar demonstrations have been given elsewhere (Brooks & Vokey, 1991; Johnstone & Shanks, 2001; Reber, Kassin, Lewis, & Cantor, 1980; Vokey & Brooks, 1992; Whittlesea & Dorken, 1993; Whittlesea & Wright, 1997; Wright & Whittlesea, 1998).

Although classification has been a popular technique for investigating implicit learning, it is a blunt instrument. The grammars that have been used dictate a sequential structure for consecutive letters and where they appear in strings. For example, most of the grammars used to construct stimuli allow strings to begin only with four bigrams (e.g., Brooks & Vokey, 1991; Reber, 1967, 1969). If a participant classifies strings on the basis of knowledge of the first two permissible letters, he or she could discriminate grammatical strings. In fact, as long as a participant adopts a decision strategy based on information that is correlated with the rules of the grammar or that captures even one of the rules, he or she will discriminate test items successfully. The correlation between the different kinds of information that participants use to classify test strings has preoccupied the field: Do participants use knowledge about legal bigrams, positional dependencies, grammatical rules, or whole string similarity to make decisions? The classification technique makes it difficult to answer the question definitively (Johnstone & Shanks, 1999, 2001).

Another weakness of using a classification task to examine implicit learning is that discrimination reflects the participants' concept of structure once they are told that the stimuli were structured according to rules. Once participants are told that the stimuli are structured, they are able to reassess what they learned and discriminate strings on the basis of information from that reassessment. The recall task used by Reber (1967) and Miller (1958), however, measures learning of grammatical structure without telling the participants that the stimuli are structured according to a grammar. Therefore, we examine implicit learning using the recall rather than the classification task.

The present article extends the arguments by Brooks and his colleagues (Brooks, 1978; Brooks & Vokey, 1991; Higham, 1997; Vokey & Brooks, 1992, 1994; Whittlesea & Dorken, 1993; Whittlesea & Wright, 1997; Wright & Whittlesea, 1998), using ideas from Perruchet and Gallego's (1997) subjective-unit account, to examine the recall advantage for grammatical sequences. We quantify how grammatical constraint produces local structure in exemplars and demonstrate that participants exploit local structure during recall.

Experiment 1

Experiment 1 was conducted to demonstrate the standard memorial advantage for grammatically constrained stimuli in a serial-recall task and to serve as a point of departure for the work to follow. Participants studied and recalled sequences of colored circles. We used sequences of colors rather than sequences of letters to avoid interference from sequential letter dependencies in printed English.

Our task honored Murdock's (1974) classic distinction between item and order information. Participants reported the sequences in the same order in which they were presented. At report, the colors used to construct sequences were provided. By forcing serial report of the colors and by providing the colors from the sequence, we directed learning toward the sequential dependencies defined by the grammar.

We used a different kind of Markov grammar than is typically used in the literature. The kinds of Markov grammar used to investigate implicit learning (called finite state grammars) impose constraints both on the position of symbols within strings and on permissible string lengths. As Johnstone and Shanks (1999) have argued, these grammars introduce complexities that make analysis of learning difficult. Hence, these authors suggested abandoning the use of Markov grammars in favor of a more tractable alternative (e.g., biconditional grammars).

We agree that using Markov grammars can make analysis difficult. However, the difficulty is not with Markov grammars themselves but with the use of overly complicated grammars. We use Markov grammars that defer positional constraints and that produce enough sequences of equal length so we can finesse difficulties of analysis (i.e., stationary Markov grammars with uniform initial state vectors).

Each stimulus comprised eight events selected from a set of six colors. The sequences were constructed according to either the *constrained* or the *control* grammar shown in Table 1. Both grammars show the probability with which each of six colors can follow one another in successive positions in a sequence (from position n in a sequence to position $n + 1$ in a sequence). The constrained grammar, shown in the left panel of Table 1, specifies that each of the six colors can be followed by two other colors. For example, Color 1 can be followed by either Color 2 or Color 3 with a probability of .5 but cannot be followed by Colors 1, 4, 5, or 6. The control grammar, shown in the right panel of Table 1, specifies that each of the six colors can follow any other with equal likelihood, subject to the constraint that no color can follow itself.

The grammars that we used solve the problems associated with finite state grammars we noted earlier. In particular, (a) they do not impose positional constraints (i.e., all six colors are equally likely to occur in each of the eight positions in a string), (b) the grammars describe the transition probabilities among the six symbols for all positions in a sequence, (c) the grammars can be used to produce enough strings of the same length to avoid strings of different lengths (e.g., $6 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 = 768$ strings of eight numbers each according to the constrained grammar), and (d) because the grammar describes only first-order symbol transitions, without introducing positional constraints and complexities based on sequence length, it is possible to count the number of grammatical violations in a string.

Method

Participants. Eighteen students from the Queen's University at Kingston psychology undergraduate participant pool participated in the study. The participants were assigned randomly to one of two treatment groups defined by the grammars in Table 1. All participants had normal or corrected-to-normal vision, and none had a visual color deficiency.

Apparatus. The experiment was administered on a personal computer equipped with a 17-in. monitor and a standard mouse. The monitor's resolution was set to $1{,}024 \times 768$ pixels. Participants interacted with the program by using the mouse to click on symbols displayed on the monitor.

Stimuli. A set of 20 sequences was generated for each participant. All sequences comprised eight events and conformed with either the constrained grammar or the control grammar (both shown in Table 1).

The constrained grammar, shown in the left panel of Table 1, specifies that each color can be followed by two of six colors with probability .5. The control grammar, shown in the right panel of Table 1, specifies that each color can be followed by any color except for itself with probability .2.

For each sequence, the first event was selected at random (1 through 6), and successive events were selected according to the transition probabilities in the relevant grammar. After eight events had been selected, a sequence was completed.

To get a set of 20 sequences, we sampled sequences one at a time according to the transition probabilities in the appropriate grammar. If a sampled sequence was not already in the stimulus set, we added it; otherwise, we discarded and replaced it. The process continued until the stimulus set comprised 20 unique sequences that conformed to the relevant grammar.

After 20 sequences had been selected, the six colors, blue (B), yellow (Y), red (R), green (G), fuchsia (F), and purple (P), were mapped randomly to the digits 1 through 6. The digit–color mapping was used to rewrite the 20 sampled digit sequences as color sequences. For example, if red, green, blue, yellow, fuchsia, and purple were assigned to the digits 1 through 6, respectively, then the sequence *12341256* was rewritten as *RGBYRGFP*. The color–digit mapping was randomized independently for each participant. The stimulus sequences were presented to participants as series of eight colored circles.

Procedure. Participants were tested individually. A session began with a preview of the six colors used to construct sequences. After participants confirmed that they could discriminate the colors, they were informed that they would study and recall sequences of colored circles.

Each trial began with the word *Study* printed on the screen. The participant initiated a trial by clicking on the word with a computer mouse. Next, a series of eight colors was presented. Each color was shown in a circle 233 pixels in diameter on the left side of the screen. At a viewing distance of approximately 70 cm, the visual angle of the circles was approximately

Table 1
The Constrained and Control Grammars Used to Construct Sequences in Experiments 1 and 2

	Constrained grammar (n + 1)						Control grammar (n + 1)					
Colors (n)	1	2	3	4	5	6	1	2	3	4	5	6
1	.0	.5	.5	.0	.0	.0	.0	.2	.2	.2	.2	.2
2	.0	.0	.5	.5	.0	.0	.2	.0	.2	.2	.2	.2
3	.0	.0	.0	.5	.5	.0	.2	.2	.0	.2	.2	.2
4	.0	.0	.0	.0	.5	.5	.2	.2	.2	.0	.2	.2
5	.5	.0	.0	.0	.0	.5	.2	.2	.2	.2	.0	.2
6	.5	.5	.0	.0	.0	.0	.2	.2	.2	.2	.2	.0

Note. The grammars show transition probabilities from position n of a sequence to position n + 1 of a sequence for six colors labeled 1 though 6 in the column and row headers.

7.1°. Each of the colors was presented for 1 s, with a blank gray screen shown for 250 ms between successive colors.

Two and a half seconds after the eighth color had been removed from the screen, a response palette appeared on the right side of the screen. At a viewing distance of approximately 70 cm, the visual angle of the response palette was approximately 5.9°. The response palette consisted of two rows of three boxes. Each of the boxes was 49 pixels square. At a viewing distance of approximately 70 cm, the visual angle of the boxes was approximately 1.7°. The six colors were assigned randomly to the six boxes on the response palette for each recall attempt.

The participants used the mouse to select the color patches to reproduce the studied sequence. They had to select a sequence of eight colors—guessing if necessary—to finish a recall attempt.

If recall was not perfect, the word *Study* reappeared on the screen, indicating to the participant that he or she was being given another opportunity to study and to recall the same sequence. When recall was perfect, a tone sounded, and a written message appeared on the screen telling the participant to move onto the next trial—that is, that he or she would study and recall a different color sequence. We measured performance by counting the number of trials that the participant took to learn each sequence.

After 20 sequences had been recalled, the participant was informed that rules had been used to construct all the 20 sequences. He or she was asked to articulate the rules and to report his or her study strategies.

Results

The data were collapsed into four blocks of five sequences each. Figure 1 shows the mean number of attempts to recall sequences correctly as a function of block and the type of grammar used to construct the sequences.

Recall of constrained sequences showed greater improvement across the four blocks (improving by 1.67 trials) compared against recall of control sequences (improved by only 0.51 trials), $F(1, 16) = 11.76$, $p < .05$, for the linear interaction. The linear interaction forced a main effect advantage for constrained (M = 2.72 trials, SE = 0.35) over control sequences (M = 3.67 trials, SE = 0.34), $F(1, 16) = 6.94$, $p < .05$. The pattern of results

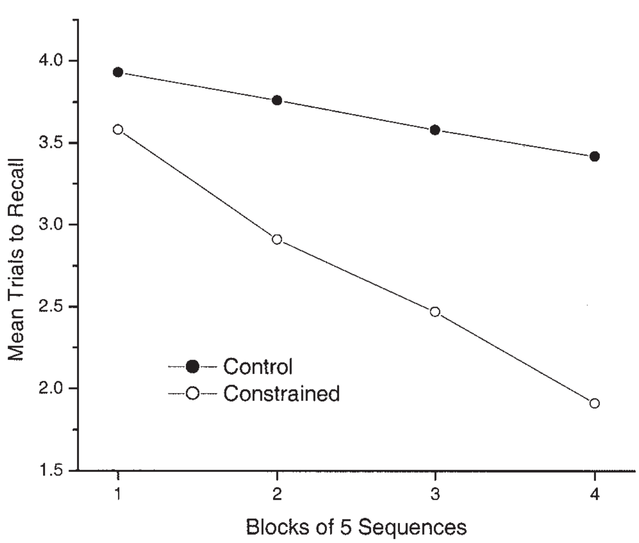

Figure 1. Mean number of trials to recall sequences as a function of block and the type of grammar (constrained or control) used to construct sequences in Experiment 1.

confirms the classic memorial advantage for constrained materials reported by Reber (1967).

Participants rarely reported knowledge of sequential dependencies. When they did so, they identified one or two color bigrams that appeared frequently in studied materials. Most of the reports described schemes for recoding the sequences to improve learning. For example, almost all participants reported recoding the sequences of eight colors into two shorter sequences of four colors. One participant wrote, "I could remember patterns like *RGBY-RGBP* by remembering *RGB-Y-RGB-P*." Participants also reported using extraexperimental associations among the colors to remember the sequences (e.g., associating colors by holiday themes).

Discussion

Experiment 1 confirms the classic advantage for grammatically constrained stimuli without participants being able to articulate the grammar. Our procedure differed in several ways from Reber's (1967) classic study. (a) Participants in our experiment studied and tried to recall only one sequence on each trial, whereas Reber's participants studied and tried to recall four sequences on each trial. (b) We used colors to construct sequences, whereas Reber used letters. (c) We presented the colors in a sequence successively (as a series), whereas Reber presented the letters of a sequence simultaneously (as a nonsense word). (d) We required participants to report the sequences in the same order in which they were presented, whereas Reber did not constrain how participants reported letter sequences. (e) We provided the set of colors used to construct sequences during recall (in the response palette), whereas Reber did not provide the letters used to construct sequences during recall. (f) Our grammar did not impose positional constraints on colors, whereas Reber's grammar did impose positional constraints on letters. In spite of these differences, we obtained the same memorial advantage for constrained materials.

Reber (1967) took the advantage for constrained sequences, without a corresponding ability to articulate the grammar, as evidence that participants knew the grammar implicitly and exploited that knowledge to recall the constrained sequences. A close look at the stimuli, however, offers an alternative explanation.

Constrained sequences had more repeating pairs of colors than control sequences. For example, *RBGYRBPF* has a pair repetition of *RB*. If participants exploited pair repetitions within individual sequences to aid recall, we can explain the advantage for constrained sequences without inferring that participants implicitly knew the grammar. The question, then, is whether structure in individual sequences, rather than grammatical structure, can account for the advantage for constrained sequences.

Garner (1974) and Miller (1953, 1956a, 1956b) have demonstrated that structure (i.e., redundancy) in an individual stimulus can be used to improve recall (Miller, Bruner, & Postman, 1954). We apply the concept to our task. To do so, we need to distinguish redundancy tied to the grammar itself—characteristics that might form the basis of implicit learning—from redundancy that is local to individual sequences. We developed tools using information theory to quantify (a) the amount of redundancy in a generative grammar (i.e., *grammatical redundancy*) and (b) the amount of redundancy in individual stimuli (i.e., *local redundancy*).

The distinction between grammatical and local redundancy derives from work by Garner (1974). A stimulus's grammatical redundancy is determined by the size of the stimulus set from which it is drawn. The smaller the set is, the greater the stimulus's grammatical redundancy is. A stimulus's local redundancy, by contrast, is determined by the number of alternative stimuli that can be generated when one reconfigures the elements of which it is composed. The complete set of alternative stimuli is called the stimulus's *inferred subset*. The smaller a stimulus's inferred subset is, the greater its local redundancy is.

Grammatical redundancy. Grammatical redundancy, G, quantifies sequential constraint in a set of stimuli, where the grammatical redundancy of a set of stimuli is at once the grammatical redundancy of each member in the set. We compute G using Shannon and Weaver's (1949) equation for uncertainty, U,

$$U = -\sum p_{ij} \log_2 p_{ij}, \tag{1}$$

where p_{ij} denotes the probability of symbol j following symbol i in a sequence (see Attneave, 1959; Garner, 1962; Miller, 1953).

We can compute the grammatical redundancy of a target grammar by comparing its uncertainty (the target grammar) against the uncertainty of an otherwise equivalent but unconstrained grammar (the reference grammar),

$$G = 1 - \frac{U\,(targetgrammar)}{U\,(referencegrammar)}. \tag{2}$$

As uncertainty of the target grammar decreases, its grammatical redundancy increases. G ranges between 0 and 1, with increasing values representing increasing grammatical redundancy.

We calculated G for both the constrained and the control grammars in Table 1. The appropriate reference grammar would allow all six colors to follow one another with equal probability (p = .16; i.e., a grammar in which there is no sequential constraint at all). G for the constrained and control grammars was about .61 and .10, respectively. The constrained grammar is more grammatically redundant than the control grammar.

By definition, all sequences generated according to the same grammar have equal grammatical redundancy. Nevertheless, some sequences from the same grammar are better structured than others. The constrained grammar in Table 1, for example, generates both *12351235* and *12345613*. Because both are generated with the same grammar, their grammatical redundancy is equal. Therefore, if grammatical redundancy controls performance, these sequences ought to be equally difficult to recall. However, one of the sequences might be better structured (or more locally redundant) than the other and, for that reason, could be easier to recall. The next section describes a way to quantify the amount of local structure in a sequence.

Local redundancy. Local redundancy of a stimulus is determined by the number of alternative stimuli that one can generate by reconfiguring elements in the original stimulus. For example, the sequence *RBG* can be reconfigured as *RGB*, *BRG*, *BGR*, *GRB*, and *GBR*—if the letters *R*, *G*, and *B* are the elements making up the stimulus. The complete set of alternative stimuli is called the stimulus's inferred subset. The smaller a stimulus's inferred subset is, the greater its local redundancy is.

Different orders of local redundancy can be computed for the same stimulus. Zero-order local redundancy, L_0, considers the number of unique symbols in a sequence but does not consider the sequential constraint among those symbols. The fewer unique symbols there are in a sequence, the greater its L_0 is. First-order local redundancy, L_1, by contrast, considers the number of unique first-order symbol transitions in a sequence (i.e., unique bigrams). The more repetitive symbol transitions in a sequence are, the greater its L_1 is.

To measure the zero-order local redundancy of a sequence, one determines the size of its inferred subset as the number of stimuli that one can generate by rearranging the individual symbols in the original. For example, the four-letter sequence *ABCD* can be reconfigured to generate 4! = 24 sequences (including the original stimulus). *ABAB*, by contrast, has two repeating letters, and, therefore, its inferred subset includes only 6 alternative sequences (including the original stimulus): {*ABAB*, *BABA*, *ABBA*, *BAAB*, *BBAA*, *AABB*}. The smaller a stimulus's inferred subset is, the greater its local redundancy is: Because *ABAB* has a smaller inferred subset, it is more locally redundant than *ABCD.*

One can quantify the zero-order local redundancy, L_0, of a sequence by comparing the number of members in its inferred subset (see the numerator in Equation 3) against the number of members in the inferred subset of a sequence of equal length that does not have any repeating symbols (the denominator in Equation 3). The formula for zero-order local redundancy is

$$L_0 = 1 - \frac{\dfrac{N!}{\Pi k_j!}}{N!}, \tag{3}$$

where N is the number of symbols in the sequence and k_j are the counts for each of the j symbols that could have occurred (i.e., all the symbols in the generative grammar). Like grammatical redundancy, L_0 ranges between 0 and 1, with increasing values representing increasing zero-order local redundancy.

One can determine the first-order local redundancy, L_1, of a sequence by rearranging the unique bigrams (i.e., first-order sequential transitions) that appear in a sequence. When first-order transitions are the unit for analysis, the inferred subset comprises all the unique sequences that can be constructed using the bigrams in the original sequence and that have the same number of elements as the original sequence. For example, *ABCD* has three unique bigrams {*AB*, *BC*, *CD*} that can be reconfigured to produce an inferred subset of 3! = 6 alternative sequences (including the original): {*ABCD, ABBC, BCAB, BCCD, CDAB*, *CDBC*}. *ABAB,* by contrast, has two unique bigrams (*AB*, *BA*) that can be reconfigured to produce an inferred subset of four alternative sequences (including the original): *ABAB*, *BABA*, *ABBA*, *BAAB*. Because *ABAB* has a smaller inferred subset than *ABCD*, it has greater first-order local redundancy.

The equation for determining first-order local redundancy, L_1, is

$$L_I = 1 - \frac{\dfrac{(N-1)!}{\Pi k_{ij}!}}{(N-1)!}, \tag{4}$$

where N is the length of the sequence and k_{ij} are counts for each of the unique ij bigrams (i.e., transitions from a symbol i to a symbol j) that could have occurred (i.e., all the bigrams in the grammar). The smaller a stimulus's inferred subset is, the greater its first-order local redundancy is. Like grammatical redundancy and zero-order local redundancy, L_1 ranges between 0 and 1, with increasing values representing increasing first-order local redundancy.

The concept of an inferred subset was developed by Garner (1974) and has been used to predict perception of, and memory for, two-dimensional visual patterns. Patterns that have smaller inferred subsets are judged to be simpler than patterns that have larger inferred subsets (Clement & Vernadoe, 1967; Garner & Clement, 1963; Garner & Whitman, 1965), and recall accuracy for a pattern varies inversely with the size of the pattern's inferred subset (Attneave, 1955; Schnore & Partington, 1967).

Armed with the quantification of local redundancy, we reexamined data from Experiment 1. The left panel of Figure 2 shows the mean number of trials to recall sequences as a function of zero-order local redundancy (L_0).

The closed circles show the relationship for the control group, and the open circles show the relationship for the constrained group. Because a different set of sequences was generated at random for each participant, the points are based on different numbers of observations. For the constrained grammar, the points are based on 71, 5, 68, 3, 21, and 12 observations for L_0 = .75, .83, .88, .92, .94, and .96, respectively. For the control group, the points are based on 27, 4, 55, 49, 7, 28, 5, 3, and 2 observations for sequences, with L_0 = .75, .83, .88, .92, .94, .96, .97, .98, and .99, respectively. In both cases, as L_0 increased, participants generally took fewer trials to recall a sequence. Performance was better overall for participants learning constrained sequences.

Because we set out to examine the influence of sequential constraint on recall, we reexamined the data from Experiment 1 as a function of first-order local redundancy. The right panel of Figure 2 shows the mean number of trials that participants took to recall sequences as a function of their first-order local redundancy (L_1).

For the control grammar, the points are based on 107, 45, 23, 4, and 1 observations for L_1 = .00, .50, .75, .88, and .94, respectively (we did not include the point for L_1 = .94 because there was only the one observation in only the control condition). For the constrained group, the points are based on 45, 77, 33, and 25 observations for sequences, with L_1 = .00, .50, .75, and .88, respectively. Again, there is a clear relationship between local redundancy and the mean number of trials that participants took to recall the sequences correctly: As L_1 increased, participants took fewer trials to recall a sequence.

Because the relationship between performance and both zero-order and first-order local redundancy was consistent and because we are interested in how sequential constraint affects learning, we use L_1 rather than L_0 as the measurement of local redundancy in subsequent experiments. We did not report other-order local redundancy (a) because higher order units, such as second-order or third-order transitions, are made up from combinations of first-order transitions and (b) because our sequences were short.

Next, we examined the relationship between grammatical and local redundancy in sequences generated according to each of the grammars in Table 1. In a Monte Carlo analysis, we sampled 20 sets of 20 sequences from the constrained grammar and 20 sets of 20 sequences from the control grammar. All the sequences were generated according to the same rules that were used to generate stimuli in Experiment 1. We calculated the first-order local redundancy for each sequence.

Mean first-order local redundancy for sequences generated with the constrained grammar (L_1 = .58) was more than double the mean first-order local redundancy for sequences generated using the control grammar (L_1 = .26). Grammatical and local redundancies are partially confounded: As grammatical redundancy increased, local redundancy increased as well.

In light of both the confound between grammatical and local redundancy and the strong relationship between local redundancy and recall, it is an open question whether the advantage for constrained materials reflects participants' implicit knowledge of the grammar (the classic interpretation favored by Reber, 1967, 1969) or whether it reflects participants' exploitation of local redundancy in individual sequences—a kind of redundancy that is incidental to the grammar.

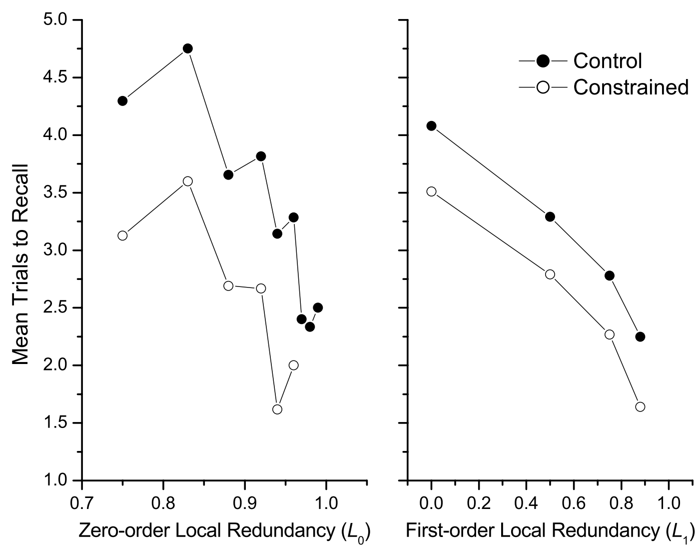

Figure 2. Mean trials to recall as a function of zero-order (L_0) and first-order (L_1) local redundancy and the type of grammar (constrained or control) used to construct sequences in Experiment 1.

Experiment 2

Experiment 2 tests whether the memorial advantage for constrained sequences persists when the first-order local redundancy of both constrained and control sequences does not differ. Because first-order local redundancy is equated in sets of constrained and control sequences, any advantage for constrained stimuli is attributable to participants' exploitation of grammatical redundancy. If learning is not better with constrained sequences, however, it would be hard to maintain the view that the participants in Experiment 1 exploited grammatical redundancy. Instead, we would argue that the benefit for constrained sequences in Experiment 1 reflects exploitation of local redundancy that derives from grammatical constraints.

Method

Participants. Twenty-four students from the Queen's University at Kingston psychology undergraduate participant pool participated in the study. The participants were assigned randomly to one of two treatment groups defined by the grammars in Table 1. All participants had normal or corrected-to-normal vision, and none had a visual color deficiency.

Apparatus and procedure. The experiment was administered with the same apparatus as in Experiment 1. The procedure was the same as in Experiment 1 except that the stimuli were generated differently.

Stimuli. A set of 20 sequences was generated for each participant. All sequences comprised eight events and were generated according to the constrained or control grammars in Table 1. The same grammar was used to construct all 20 sequences in a set. The sequences were constructed in the same way described in Experiment 1.

The constrained and the control grammar were each used to construct six sets of 20 sequences. The first-order local redundancy of sequences in the sets was measured, and then sequences were sampled from the complementary grammar (the control grammar if the set comprised constrained sequences, and vice versa) until the first-order local redundancy of the 20 sampled sequences matched the first-order local redundancy of sequences in the yoked set (i.e., some of the sequences were rejected and replaced until their local redundancy matched that of the yoked sequence).

After 20 sequences had been selected to the stimulus sets, the six colors were mapped randomly to the digits 1 through 6. The digit–color mapping was used to rewrite the 20 sampled digit sequences as color sequences and was random for each participant.

Results

The data were arranged to reflect participants' performance on four blocks of five sequences. Figure 3 shows the mean number of study–test trials that participants took to recall sequences as a function of block and the type of grammar used to construct the sequences.

In contrast to Experiment 1, there was no interaction between sequence type and blocks and, therefore, no evidence of a memorial advantage for grammatically constrained sequences, $F(3, 66) < 1.00$. When we averaged across performance with constrained and control sequences, recall improved by a mean 0.81 trials across the four blocks, $F(3, 66) = 6.82$, $p < .05$, with the linear trend accounting for 84.6% of the variance, $F(1, 22) = 16.28$, $p < .05$.

Figure 4 shows the mean number of trials to recall a sequence depending on its first-order local redundancy. For both the constrained and the control group, points are based on 102, 93, 33, and 12 observations for sequences with $L_1 = .00$, .50, .75, and .88,

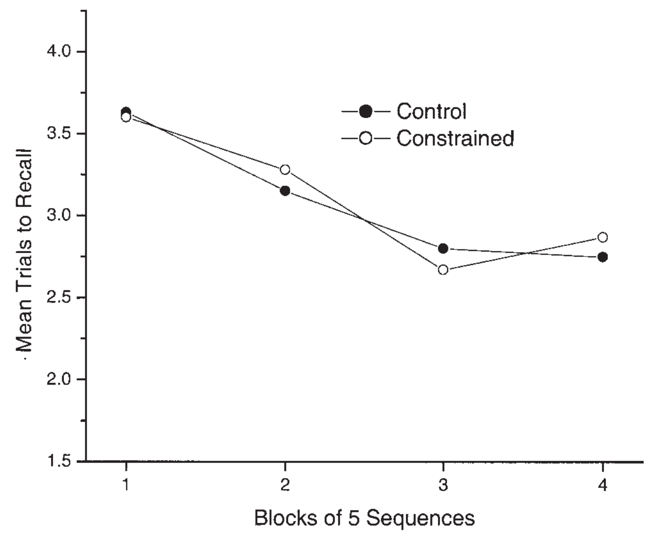

Figure 3. Mean trials to recall as a function of block and the type of grammar (constrained or control) used to construct sequences in Experiment 2.

respectively. Performance improved as local redundancy increased. As in Experiment 1, participants took fewer trials to recall sequences as first-order local redundancy increased.

One might object to our conclusion on the grounds that the null difference between constrained and control sequences reflects the way that we selected the stimuli. To verify that the difference in grammatical redundancy of constrained and control sequences was not compromised, we computed the grammatical redundancy for the stimulus sets presented to participants in Experiment 1 and in Experiment 2. We then compared mean grammatical redundancy in the constrained and control conditions across the two experiments. Mean grammatical redundancy was the same in Experiments 1 and 2 both for sets of constrained sequences ($G = .62$) and for sets of control sequences ($G = .15$). The way we selected the stimuli did not change the grammatical redundancy in sets of constrained and in sets of control sequences from Experiment 1. Therefore, the null difference between constrained and control sequences does not reflect the difference between Experiments 1 and 2 in how the stimuli were selected.

Participants rarely reported knowledge of sequential dependencies. In cases where they did so, they identified one or two color bigrams that appeared frequently in studied materials. Most of the reports described schemes for recoding the sequences to improve learning (e.g., recoding the sequences of eight colors into two shorter sequences of four colors). Participants used extraexperimental associations among the colors to remember the sequences.

Discussion

When we equated the first-order local redundancy in sets of constrained and control sequences but allowed the grammatical redundancy of constrained and control sequences to differ, we found no evidence of a memorial advantage for constrained over control sequences. The null difference makes it difficult to maintain the argument that participants exploited grammatical redun-

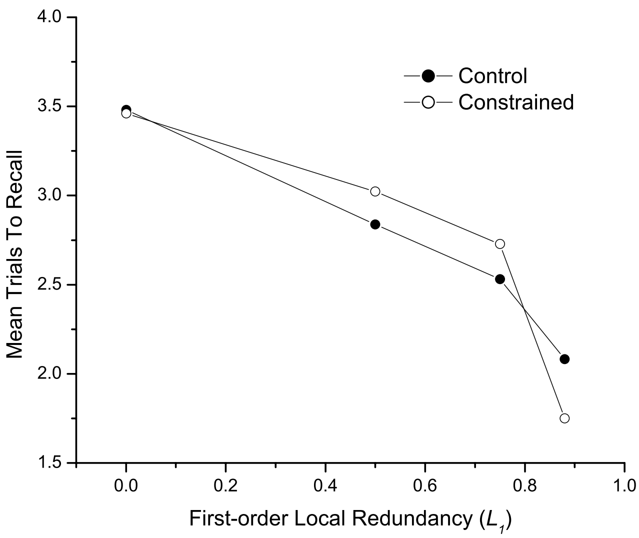

Figure 4. Mean trials to recall as a function of first-order local redundancy (L_1) and the type of grammar (constrained or control) used to construct sequences in Experiment 2.

dancy in Experiment 1. Instead, we suggest that learning in Experiment 1 benefited from participants' use of first-order local redundancy that was incidental to grammatical redundancy.

Our suggestion is, of course, consistent with the results of both experiments, but it is in conflict with the claim that participants knew the grammar. We do not need to infer that participants knew the grammar (either implicitly or explicitly) to explain performance.

It would be misleading to conclude, however, that learning is driven by stimulus properties alone. The way a subject organizes the stimulus has profound consequences for how well he or she can recall it at a later time (e.g., Mandler, 1967; Tulving, 1962; Tulving & Pearlstone, 1966; Tulving & Thomson, 1973; Whittlesea & Dorken, 1993). For example, *RGBYRGBP* and *RGBYPRGB* both have first-order local redundancy equal to .75, but they may not be equally difficult to learn. The way participants actively encode the sequences might make one sequence easier to recall than the other. If, for example, both sequences are encoded as two groups of four colors, then *RGBYRGBP* might be easier to learn than *RGBYPRGB* because *RGB* begins each of the encoded chunks in the first string but not in the second.

Unfortunately, it is easy to speculate on how participants encode sequences, but it is difficult to quantify the concept. Certainly, neither Experiment 1 nor Experiment 2 addresses how participants' encoding strategies affected their recall performance.

To examine the influence of encoding on recall performance, we adapted our concepts of redundancy and developed a measure for how difficult it is to recall a sequence on the basis of how it is encoded. Our method borrows heavily from work by G. A. Miller (1956a, 1956b, 1958) and others, notably Mandler (1967) and Tulving (1962).

Miller (1956a, 1956b) argued that recoding elements of a stimulus into higher order units can reduce the load that the original stimulus places on memory. In a now famous anecdote, he described watching engineers reading a row of 15 lights. At first, he was amazed at the speed with which they could decode and remember the row—there were too many lights for comfort. He soon discovered, however, that the engineers had grouped the lights into successive triplets and reexpressed each triplet as an octal digit. For example, the engineers reduced the memory load imposed by a sequence of lights 011000101001111 (where 0 means *off* and 1 means *on*) by reexpressing the sequence as five triplets of binary numbers, 011 000 101 001 111, and then as five octal digits, 3 0 5 1 7. To recall the original sequence, they translated the octal digits back into their binary equivalents. Our measurement for how encoding influences recall, which we call organizational redundancy, is based on Miller's (1956a, 1956b) idea.

Organizational redundancy, O, quantifies the reduction of information in a stimulus after it has been organized in a particular way. The measure indicates how well the different units in the recoded stimulus predict one another's contents. For example, recoding a sequence *RBGYRBGY* into two groups of four colors reduces the information in the sequence by one half, because the first chunk, *RBGY*, is a perfect predictor of the second chunk, *RBGY*.

We express the grouping of a sequence along two dimensions: (a) chunking units, c_i, and (b) serial positions within chunking units, p_j. A sequence is represented by a two-dimensional array, with rows representing chunking units and columns representing serial positions within the chunking units. The number of chunking units and the number of serial positions within chunking units is determined by how a sequence has been encoded. Table 2 shows two sequences, *RGBYRGBP* and *RGBYPRGB*, organized into two chunking units of four colors.

Organizational redundancy, O, for a sequence is computed from the array representation as

$$O = 1 - \frac{\sum_{j=1}^{p} \Delta_j}{N}, \qquad (5)$$

where p is the number of serial positions within chunking units, Δ_j is the number of differing colors in each of the j serial positions of the array (i.e., the columns of the array), and N is the length of the sequence. O ranges from 0 to 1, with increasing values representing increasing organizational redundancy.

Table 2
Notation for the Sequences RGBYRGBP and RGBYPRGB Based on How Participants Reported Organizing Sequences in Experiments 1 and 2

	Position							
Chunking unit	p_1	p_2	p_3	p_4	p_1	p_2	p_3	p_4
c_1	R	G	B	Y	R	G	B	Y
c_2	R	G	B	P	P	R	G	B
Δ_j	1	1	1	2	2	2	2	2

Note. The sequences are organized along two dimensions: chunking units, c_i, and positions within chunks, p_j. The row labeled Δ_j indicates the number of differing symbols at each of the j serial positions across chunking units.

The first-order local redundancy of the sequences in Table 2, *RGBYRGBP* and *RGBYPRGB*, suggests that the sequences will be equally difficult to recall (both have a first-order local redundancy of .75). The organizational redundancy of the two sequences, by contrast, suggests that *RGBYRGBP* (O = .375) will be easier to recall than *RGBRGBYP* (O = .000): The two kinds of redundancy are confounded.

We reexamined recall performance in Experiments 1 and 2 as a function of organizational redundancy, given that sequences were grouped into two groups of four colors (as our participants indicated that they had done). Figure 5 shows the relationship between organizational redundancy and the mean number of trials to recall sequences for Experiments 1 (closed circles) and 2 (open circles). Recall was strongly related to organizational redundancy.

Next, we examined the confound between first-order local redundancy and organizational redundancy. In a Monte Carlo analysis, we sampled 250 sequences each for L_1 = .00, .50, .75, .88, and .99. All sequences were selected randomly according to the control grammar shown in Table 1. The same rules were used to generate stimuli as in Experiment 1. The mean organizational redundancy for sequences with first-order local redundancy equal to .00, .50, .75, .88, and .99 was .07, .09, .14, .24, and .30, respectively. Local and organizational redundancies are partially confounded: As local redundancy increased, organizational redundancy increased as well.

In light of both the confound between local and organizational redundancy and the strong relation between organizational redundancy and recall in Experiments 1 and 2, it is unclear whether the advantage for grammatically constrained materials reflects participants' use of local redundancy (incidental to the grammatical redundancy) or their use of organizational redundancy (incidental to local redundancy and therefore by association also incidental to grammatical redundancy).

Experiments 3 and 4 were designed to examine the chain of relationships among organizational, local, and grammatical redundancy. Because organizational redundancy is based on how participants encode sequences, we introduced a pause during the presentation of sequences to encourage an encoding consistent with how our participants in Experiments 1 and 2 described organizing the sequences. Experiment 4 uses the same technique to show that the way a sequence is encoded can both benefit and cost recall.

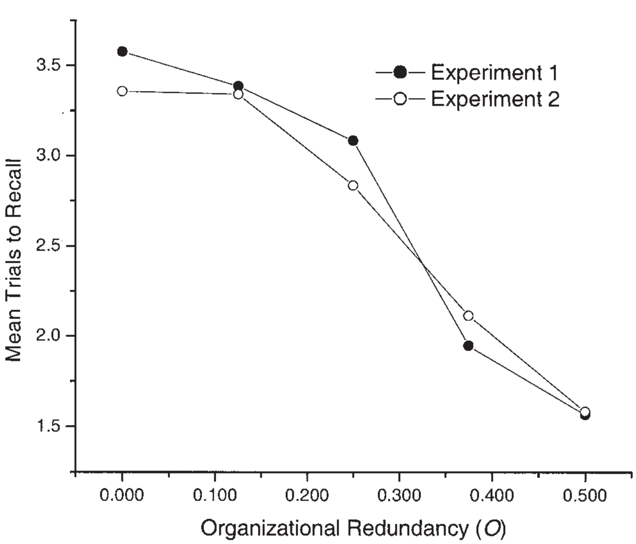

Figure 5. Mean trials to recall sequences as a function of organizational redundancy (O) in Experiments 1 and 2.

Experiment 3

In Experiment 3, we unconfounded first-order local redundancy from organizational redundancy to examine how the two predict performance. The procedure was the same as in Experiment 1, with one exception. A 1-s pause was inserted between the presentation of the fourth and fifth colored circles of a sequence to encourage participants to group the series of eight colors into two units of four colors. The sequences that were tested in the experiment had first-order local redundancy equal to either .75 or .50 and had organizational redundancy equal to or greater than zero.

Method

Participants. Twenty students from the Queen's University at Kingston psychology undergraduate participant pool took part in the study. All participants had normal or corrected-to-normal vision, and none had a visual color deficiency.

Apparatus and procedure. The experiment was administered with the same apparatus as in the first two experiments. The procedure was the same as in Experiment 1, with two exceptions: A 1-s pause was inserted between the presentation of the fourth and fifth colored circles of a sequence, and participants learned 16 sequences.

Stimuli. A set of 16 color sequences was generated for each participant. The set of 16 comprised two sequences based on each of eight stimulus patterns. Sequences based on the patterns *12341235* and *12345234* had L_1 = .75 and O = .375. Sequences based on the patterns *12345123* and *12342345* had L_1 = .75 and O = .000. Sequences based on the patterns *12341256* and *12345634* had L_1 = .50 and O = .250. Sequences based on the patterns *12345346* and *12342356* had L_1 = .50 and O = .000.

We rewrote each of the stimulus patterns as color sequences by randomly mapping the six colors to the digits and then using the color–digit mapping to rewrite the sequence. The color–digit mapping was freshly randomized before a pattern was rewritten as a color sequence.

The 16 color sequences were presented in a random order, under the constraint that a sequence constructed on the basis of each of the eight stimulus patterns was tested once as one of the first 8 sequences and once as one of the second 8 sequences of the experiment.

Results

Table 3 shows the mean number of study–test trials to recall sequences correctly as a function of both first-order local redundancy and organizational redundancy. For the analysis, organizational redundancy was treated as a dichotomous variable, either equal to zero or greater than zero.

Sequences with organizational redundancy greater than zero (1.87 trials) were recalled in fewer trials than sequences with organizational redundancy equal to zero (3.46 trials), $F(1, 19)$ = 60.94, $p < .05$. Sequences with L_1 = .75 (2.51 trials) were recalled in fewer trials than those with L_1 = .50 (2.82 trials), although the trend was not reliable, $F(1, 19)$ = 2.29. In short, learning did not benefit from local redundancy unless a participant encoded a sequence to increase its organizational redundancy.

Table 3
Mean Trials to Recall Sequences as a Function of First-Order Local and Organizational Redundancy in Experiment 3

	First-order local redundancy			
	.75		.50	
	Organizational redundancy			
	.375	.00	.25	.00
Stimulus	1234.1235	1234.5123	1234.1256	1234.5346
Patterns	1234.5234	1234.2345	1234.5634	1234.2356
M	1.77	3.25	1.96	3.67
SE	0.10	0.21	0.14	0.36

Note. Periods between numbers in stimulus patterns indicate the point at which the 1-s pause was inserted during presentation.

Participants reported recoding the sequences of eight colors into two shorter sequences of four colors (as we hoped they would to be consistent with our measure of organizational redundancy). They also reported using extraexperimental associations among colors to remember the sequences.

Discussion

Recall performance was strongly related to our measurement of organizational redundancy, which was, in turn, based on how participants reported encoding sequences in Experiments 1 and 2.

The fact that organizational redundancy is determined by the way participants encode stimulus sequences complicates the story. The advantage for grammatical materials forms a chain of relationships from grammatical redundancy to local redundancy and from local redundancy to organizational redundancy. As one moves down the chain, the explanation of performance becomes more detailed and is closer to participants' reported strategies.

Experiment 4

Our analysis of organizational redundancy in Experiment 3 is limited, however, because we only calculated it in one way: on the basis of grouping sequences into two units of four colors. Experiment 4 was designed to examine the flexibility of encoding and to examine whether our measure of organizational redundancy captures that flexibility.

Table 4 shows two sequences, *RGBYRGBP* and *RGBRGBYP*, grouped in two different ways. The top panel shows the sequences grouped into two units of four colors—the organization from Experiment 3. The bottom panel shows the sequences grouped as a triplet followed by a quintuplet.

The measure of organizational redundancy predicts that *RGBYRGBP* ought to be easier than *RGBRGBYP* when the 1-s pause is inserted between the fourth and fifth colored circles of the sequence (O = .375 and .000, respectively), whereas the reverse will be true if the pause occurs between the presentation of the third and fourth colored circles (O = .000 and .375, respectively). If this is true, the participant's use of local redundancy in the sequence predicts performance, rather than tacit appreciation of local redundancy.

Participants studied and recalled series of eight colored circles. A 1-s pause was inserted between either the fourth and fifth colored circles or the third and fourth circles in the series. We calculated organizational redundancy depending on where the pause was inserted during presentation.

Method

Participants. Eighteen students from the Queen's University at Kingston psychology undergraduate participant pool took part in the study. All participants had normal or corrected-to-normal vision, and none had a visual color deficiency.

Apparatus and procedure. The experiment was administered with the same apparatus used in Experiments 1, 2, and 3. The procedure was the same as in Experiment 3, with one exception. The 1-s pause was inserted either between the fourth and fifth events of a sequence or between the third and fourth events of a sequence.

Stimuli. A set of 16 sequences was generated for each participant with four stimulus patterns (*12341235*, *12312345*, *12341256*, *12312456*). Each of the stimulus patterns was used to construct 4 sequences. Sequences based on patterns *12341235* and *12312345* had L_1 = .75. Sequences based on patterns *12341256* and *12312456* had L_1 = .50.

We rewrote each of the stimulus patterns as a color sequence by randomly mapping the six colors to the digits and then using the color–digit mapping to rewrite the sequence. The color–digit mapping was freshly randomized before a pattern was rewritten as a color sequence.

Each of four stimulus patterns was tested twice as one of the first eight sequences in the experiment (once with a pause introduced between the fourth and fifth colored circles of a sequence, and once with the 1-s pause introduced between the third and fourth colored circles of a sequence) and twice as one of the second eight sequences of the experiment (once with a pause introduced between the fourth and fifth colored circles of a sequence, and once with the 1-s pause introduced between the third and fourth colored circles of a sequence).

Calculation of organizational redundancy depended on where the 1-s pause was inserted during a sequence's presentation. When the pause was

Table 4
Notation for Two Sequences, RGBYRGBP and RGBRGBYP, Based on a Simple Chunking Model That Describes Organization Along Two Dimensions: Chunking Units, c_i and Positions Within Chunks p_j

Chunking unit	Position									
	p_1	p_2	p_3	p_4	p_5	p_1	p_2	p_3	p_4	p_5
	Four colors and four colors									
c_1	R	G	B	Y		R	G	B	R	
c_2	R	G	B	P		G	B	Y	P	
Δ_j	1	1	1	2		2	2	2	2	
	Three colors and five colors									
c_1	R	G	B			R	G	B		
c_2	Y	R	G	B	P	R	G	B	Y	P
Δ_j	2	2	2	1	1	1	1	1	1	1

Note. The top panel shows the two sequences organized into two units of four colors. The bottom panel shows the two sequences organized into a triplet comprising the first three colors of a sequence and a quintuplet comprising the last five colors of a sequence. The row labeled Δ_j indicates the number of differing symbols at each of the j serial positions of the array.

inserted between the fourth and fifth colored circles of a sequence, *1234.5678* (the period signifies the 1-s pause), organizational redundancy was calculated as if the sequence were encoded as two separate units of four colors. When the pause was introduced between the third and fourth colored circles (i.e., *123.45678*) organizational redundancy was calculated as if the sequence was encoded as a triplet and a quintuplet of colors.

Table 5 shows the stimulus patterns and their first-order local and organizational redundancy (on the basis of when the 1-s pause was introduced into a sequence). Because organizational redundancy depends on the placement of the pause, each of the four patterns took two values of organizational redundancy, depending on when the pause was inserted.

Results

Table 5 shows the mean number of trials to recall sequences correctly as a function of first-order local redundancy and organizational redundancy. For the analysis, organizational redundancy was treated as a dichotomous variable: equal to zero or greater than zero.

Sequences with organizational redundancy greater than zero (1.88 trials) were recalled in fewer trials than sequences with organizational redundancy equal to zero (3.43 trials), $F(1, 17) = 72.45$, $p < .05$. Sequences with $L_1 = .75$ (2.47 trials) were recalled in fewer trials than those with $L_1 = .50$ (2.85 trials), but the difference was marginal, $F(1, 17) = 4.34$, $p = .05$. Consistent with Experiment 3, learning did not benefit from local redundancy unless a participant encoded a sequence to increase its organizational redundancy.

Figure 6 shows trials to recall sequences as a function of organizational redundancy. Trials to recall decreased linearly as organizational redundancy increased, $F(1, 17) = 78.76$, $p < .05$; the linear trend accounted for 99.18% of the variance.

Most participants reported using the 1-s pause to rehearse the sequence and often mentioned using grouping strategies to remember the sequences. Several participants reported developing and then using ideas about similarities and distinctions among colors to improve learning (e.g., remembering colors as if they were combinations of one another).

Discussion

Although we cannot guarantee that participants encoded sequences on the basis of when a pause was inserted on each trial, the nearly perfect linear relationship between organizational redundancy and recall suggests that they did so (at least on average). The relationship between organizational redundancy and the number of trials to recall sequences provides converging evidence for our conclusion from Experiment 3. Performance depended on how participants grouped sequences, and local redundancy did not benefit performance independently of its relationship with organizational redundancy.

Table 5
Mean Trials to Recall Sequences as a Function of First-Order Local and Organizational Redundancy in Experiment 4

	First-order local redundancy .75		.50	
	Organizational redundancy .375	.00	.25	.00
Stimulus	1234.1235	123.41235	1234.1256	123.41256
Patterns	123.12345	1231.2345	123.12456	1231.2456
M	1.60	3.33	2.17	3.53
SE	0.12	0.33	0.19	0.29

Note. Periods between numbers in stimulus patterns indicate the point at which the 1-s was inserted during presentation.

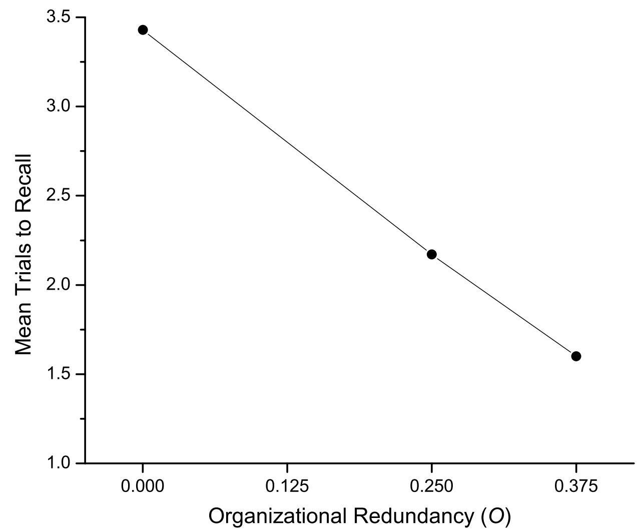

Figure 6. Mean trials to recall as a function of organizational redundancy (*O*) in Experiment 4.

Because organizational redundancy is confounded with local redundancy and because local redundancy, in turn, is confounded with grammatical redundancy, the advantage for grammatical materials in Experiment 1 can be explained by a detailed analysis of how the sequences were grouped and encoded to facilitate recall. Although participants' written reports reliably failed to show knowledge of the grammar, they have faithfully reported how they exploited grammatical constraints by how they grouped and recoded the sequences to improve memory.

General Discussion

There is a memorial advantage for sequences of letters constructed according to a grammar (Miller, 1958; Reber, 1967). Because participants are rarely able to articulate the grammar, the advantage has been taken by some as evidence that participants learn the grammar implicitly (Dienes et al., 1991; Knowlton & Squire, 1994; Mathews et al., 1989; Manza & Reber, 1997; Reber, 1989, 1993).

Experiment 1 confirms the classic advantage for constrained stimuli. We used concepts from information theory to quantify grammatical redundancy (structure in a set of stimuli) and local redundancy (structure in an individual stimulus). A Monte Carlo analysis confirmed that the two are partially confounded: As grammatical constraint increased, local redundancy increased. We reexamined performance in Experiment 1 in light of the confound. Recall performance improved as local redundancy increased. Because local redundancy can be exploited to improve recall (Garner, 1962, 1974; Garner & Degerman, 1967; Miller, 1956b, 1958) and because constrained stimuli have greater local redundancy, we argued that there should be an advantage for constrained stimuli even if participants do not know the grammar.

Experiment 2 tested whether the advantage for constrained sequences depended on the confound with local redundancy. When local redundancy was equated in sets of constrained and control sequences, we found no evidence of an advantage for constrained materials. We concluded that subjects used local redundancy and not grammatical redundancy to aid learning.

Next, we examined how participants used local redundancy to improve learning. We represented how a sequence was encoded in a two-dimensional array and developed a measure of redundancy based on that representation, organizational redundancy. A Monte Carlo analysis confirmed that local and organizational redundancies were partially confounded in Experiments 1 and 2: As local redundancy increased, organizational redundancy increased. In light of the confound, we reexamined performance in the first two experiments and found that performance improved as organizational redundancy increased.

Experiment 3 demonstrated that organizational redundancy was strongly related to recall performance. Experiment 4 corroborated the result while (a) documenting a strong linear relationship between recall and organizational redundancy and (b) showing that grouping stimuli can lead to both benefits and costs in performance. Participants showed no tacit appreciation of local redundancy, but they were able to use the affordances (Gibson, 1979) that local redundancy provided and group elements of a sequence to aid learning.

Contrary to claims for automatic abstraction of structure, we found no evidence for tacit appreciation of either grammatical or local redundancy. The strong relationship between organizational redundancy and performance indicates that participants had to actively recode the structure in a stimulus sequence to benefit from its presence.

Classification Tasks

Although we framed our ideas using a serial-recall procedure, they could be extended to explain performance in classification experiments in which participants discriminate grammatical strings after studying grammatical exemplars. Our position is closest to the processing theories that argue that the way a stimulus is encoded affects classification.

Although we agree that the way a participant encodes a stimulus has a role in discrimination, there is a need for clarity regarding how different ways of encoding produce changes in classification. If the mental transformation of stimuli at study is a critical part of an explanation, we need to describe both its contents and how that transformation takes place.

We have specified a solution by developing a formal and flexible scheme to characterize encoding variability (i.e., our measure of organizational redundancy). We expect classification to vary depending on the match–mismatch in how elements of a sequence are grouped at both study and test. For example, a sequence of colors, *RGBYRGBP*, might be endorsed as grammatical if it is grouped as *RGBY.RGBP* at both study and test. However, it may not be endorsed as grammatical if it is grouped as *RGBY.RGBP* at study and then grouped as *RGB.YRGBP* at test (i.e., transfer appropriate processing). The distinction is at the core of Whittlesea and Dorken's (1993) episodic-processing theory of learning.

Furthermore, we might expect a participant to endorse a string as grammatical if it fits a retrieval structure that had aided recall in the study phase. For example, grouping sequences of letters such as *MXVYMXVT* into two groups of four letters at study (making the string easier to remember) might encourage a participant to endorse a sequence of letters *RDCWRDCP* as grammatical because it is easier to remember when grouped in the same way. That is, strings could be endorsed as grammatical on the basis of a confluence of organization, even when they have no literal similarity. Brooks and his colleagues (cf. Brooks & Vokey, 1991; Whittlesea & Dorken, 1993) have made similar arguments.

Serial Reaction-Time Tasks

Serial reaction-time tasks represent yet another method to examine the learning of sequential structure. In these tasks, participants are seated in front of a computer that is equipped with a series of response buttons that are mapped to positions on its screen. When a light appears on the screen, the participant presses the corresponding response button. Once a response is made, the screen is reset, and a light appears at another location. The sequence continues ad nauseum.

When a grammar is used to decide the sequential structure for the lights, participants learn to respond quickly and accurately. The more constrained the grammar is, the faster that learning takes place. When participants cannot articulate the sequential constraints imposed by the grammar, it is argued that knowledge of the sequential constraints is implicit.

An alternative explanation is that participants group small runs within the continuous sequence while performing the task (Stadler, 1992, 1995; Stadler & Neely, 1997). After they have done so, the groupings stored in memory drive expectancies for what will come next. Because participants can anticipate some of the transitions, the amount of uncertainty in the display is reduced, and participants can respond more quickly.

Even a vague expectancy would be sufficient to improve performance. For example, knowing that a pattern bounces back and forth between two clusters of locations on the screen (e.g., between the left and right side of the screen) reduces the amount of information in the display by one half (or, in the terms of information theory, by one bit). Because grammatical sequences have consistent sequential dependencies and because random ones do not, an advantage for grammatical sequences is found (Perruchet, Bigand, & Benoit-Gonin, 1997; Stadler, 1992, 1995).

Our argument is consistent with our broader perspective: Gains in performance in a structured stimulus field reflect a process of organization. We expect a relationship to emerge between the organizational redundancy based on a participant's organization and reduction in the corresponding response time.

Amnesia

Evidence that is often cited to support a division of explicit and implicit memory is that performance on explicit memory tasks is impaired in cases of amnesia, whereas performance on implicit tasks is near normal. For example, amnesiacs perform almost as well as controls in discriminating grammatical from ungrammatical strings after studying grammatical exemplars, yet they are unable to recognize strings from the studied set (Knowlton &

Squire, 1992, 1994, 1996). The dissociation has been taken as prima facie evidence that implicit and explicit memory are separate and that explicit memory is impaired by the disease, whereas implicit memory is not.

The argument reminds us of the debate on whether there are separate episodic (memory for personal experiences tied to a particular time and place) and semantic (memory for meanings and concepts that are not tied to a time and place) memory stores. Tulving (1985) argued that dissociations in memory performance (e.g., recognition failure of recallable words, retention of semantic memory and loss of episodic memory in amnesia) evinced separate episodic and semantic stores. In rebuttal, Hintzman (1984, 1986) developed MINERVA 2, a resonance theory of memory, to demonstrate that semantic information can be derived at retrieval from a single store of episodic knowledge. The model has since been used by several others to describe how general knowledge can be acquired from a store of instances pertaining to a variety of contexts: acoustic abstraction (Goldinger, 1998), grammar learning (Jamieson & Mewhort, 2004; Vokey & Brooks, 1992), symbol recognition (Jones, 2001), and spelling to sound mapping (Kwantes & Mewhort, 1999).

Kinder and Shanks (2001, 2003) made a similar argument against the evidence for separate implicit and explicit systems (cf. Hintzman's, 1986, MINERVA). They used a single-store connectionist model of memory and simulated amnesia as a general encoding deficit. Although the model has only a single memory store, it obtained dissociations between performance in implicit (e.g., grammar learning, semantic priming) and explicit (e.g., recognition) memory tasks. On the basis of the success of their simulations, Kinder and Shanks argued that a single store of knowledge underlies performance on implicit and explicit tests and that the dissociation reflects different retrieval processes that are recruited for the different tests. Shanks and Perruchet (2002) gave a similar demonstration to explain dissociations between learning a sequentially constrained sequence and an inability to recognize the sequence (i.e., the dissociation between performance and recognition in serial reaction-time tasks).

Others have questioned the validity of dissociation logic as a technique for establishing separate memory stores (Neal & Hesketh, 1997; Shanks & St. John, 1994; St. John & Shanks, 1997). Dunn and Kirsner (1988), for example, demonstrated that it is possible to obtain a dissociation on two tasks with only a single system and argued that a different method must be adopted to firm up claims for separate systems. They provided an alternative method called reversed association to replace dissociation logic (although the technique has not been adopted widely). Higham, Vokey, and Pritchard (2000; see also Higham & Vokey, 2000, and Vokey & Higham, 2004) also argued against dissociation logic and provided yet another alternative technique to replace dissociation logic: opposition logic, which is a variant of Jacoby's (1991) process dissociation procedure.

We are not inclined to accept the dual-system arguments until the work by Kinder and Shanks (2001, 2003) is discredited and until reversed association (Dunn & Kirsner, 1988) or the opposition logic (Higham et al., 2000) technique are used to bolster the evidence that currently is reliant on the validity of dissociation logic.

Changed Symbol-Set Performance

Reber (1969) showed that changing the symbol set used to instantiate strings does not affect participants' capacity to learn grammatical sequences. Changing the grammar used to construct strings, however, does make it more difficult to recall grammatical stimuli. The result is taken as strong evidence that subjects must learn the abstract grammar. Although we did not examine the consequences of changing either the grammar or the symbols used to construct strings, we do not find the result troublesome to our account.

If participants develop a set of retrieval strategies that aid in exploiting local structure in studied sequences (e.g., grouping sequences of eight symbols into two shorter groups of four), changing the symbols (e.g., from letters to tones) will not affect performance: The same strategies will work no matter what symbols are used to instantiate the sequences.

By contrast, when a different grammar is suddenly used to construct stimuli, the retrieval structures that were being used to aid learning (i.e., to increase organizational redundancy in sequences) are no longer valid. They must be abandoned, and new ones must be developed. The time that it takes to develop and then apply those new strategies leads to costs in performance.

The Concept of an Inferred Subset

In developing our measure of local redundancy, we reintroduced Garner's (1974) concept of an inferred subset. We were cautious not to reify the concept as a mental operation. Garner, however, was bolder and illustrated his point with an elegant example.

He showed participants a stimulus composed of two concentric circles that were centered in a frame and asked them to describe it. Most described it as a double circle. He then showed the same stimulus along with a pair of larger concentric circles and asked participants to describe the first stimulus. Most of the participants now included the size of the circles in their description. Finally, he showed both the original stimulus and the second pair of circles along with a third pair of concentric circles located at the left edge of the frame. Most respondents now included the spatial location of the circle in their description.

Garner's (1974) point was that a single stimulus has no meaning without a context of alternatives. Using the term *circle* to describe the first stimulus implies that the concentric circles could have taken some other form, a square or triangle, perhaps. Describing the circle as having two lines implies that it might have had fewer or more lines. Describing a pair of circles as small implies that they might have some other size. Finally, describing a pair of circles as centered implies that they might have been off center.

In short, a stimulus is defined by its relation to stimuli that it might have been. The inferred subset comprises a listing of stimuli that it could have been, and we have demonstrated that the size of an inferred subset can predict learning. Others have demonstrated that the size of a stimulus's inferred subset predicts equally interesting phenomena (e.g., pattern goodness). Whether the concept describes a mental operation, however, remains an open question.

Other Grammars

We showed a confound from grammatical to local redundancy and from local to organizational redundancy with the grammars shown in Table 1. But does our case generalize to other grammars?

We examined several grammars and discovered that the confound between first-order local redundancy and grammatical redundancy varies in severity depending on the grammar and how letters are associated with transitions between nodes (e.g., if letters consistently follow one another in the grammar, the confound becomes more severe, but if they are followed by different letters, the confound is alleviated). The trouble is that the severity of the confound varies when the letters associated with grammatical transitions change. Therefore, the confound must be assessed for particular applications of particular grammars. Although we cannot make a general statement about the confound related to published work, a situation in which grammatical and local structure are not correlated would be surprising, as structure in a set of stimuli (grammatical redundancy) is a product of structure in the individual stimuli (local redundancy) of which the set is composed.

Summary

In studies of implicit learning, experimenters have focused on the relationship between learning and the grammar used to construct sequences. Consequently, the chain of relationships from grammatical to local redundancy and from local redundancy to organizational redundancy has been largely overlooked. Tracing performance through that chain, however, provides a clear explanation for the advantage for grammatical stimuli that is consistent with how participants reported trying to remember the sequences. Although we are not the first to note that subjects use local structure to aid learning, we have provided a framework to characterize and measure structure on the basis of how subjects organize sequences of symbols.

Using Garner's (1974) concept of an inferred subset and Miller's (1956a, 1956b) unitization hypothesis, we have developed measures of structure in a set of stimuli, in individual stimuli, and in stimuli as encoded. We used those measures, along with subjects' verbal reports, to show how subjects actively exploited local expressions of grammatical structure by recoding stimuli so that they are easier to retrieve. We found no evidence for tacit appreciation of either grammatical or local structure in stimuli.

Because local redundancy is correlated with grammatical redundancy, using local redundancy to organize a sequence is an indirect exploitation of grammatical redundancy. We deny, however, that the memorial advantage for grammatical stimuli is evidence that subjects know the generative grammar.

References

Attneave, F. (1955). Symmetry, information, and memory for patterns. *American Journal of Psychology, 68,* 209–222.

Attneave, F. (1959). *Applications of information theory to psychology: A summary of basic concepts, methods, and results*. New York: Holt, Rinehart, and Winston.

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–211). Hillsdale, NJ: Erlbaum.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General, 120,* 316–320.

Clement, D. E., & Vernadoe, K. W. (1967). Pattern uncertainty and the discrimination of visual patterns. *Perception & Psychophysics, 2,* 427–431.

Dienes, Z., Broadbent, D. E., & Berry, D. C. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 875–887.

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General, 113,* 541–555.

Dunn, J. C., & Kirsner, K. (1988). Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review, 95,* 91–101.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.

Garner, W. R. (1974). *The processing of information and structure*. New York: Wiley.

Garner, W. R., & Clement, D. E. (1963). Goodness of pattern and pattern uncertainty. *Journal of Verbal Learning and Verbal Behavior, 2,* 446–452.

Garner, W. R., & Degerman, R. L. (1967). Transfer in free-recall learning of overlapping lists of nonsense words. *Journal of Verbal Learning and Verbal Behavior, 6,* 922–927.

Garner, W. R., & Whitman, J. R. (1965). Form and amount of internal structure as factors in free-recall learning of nonsense words. *Journal of Verbal Learning and Verbal Behavior, 4,* 257–266.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105,* 251–279.

Gomez, R. L., & Schvaneveldt, R. W. (1994). What is learned from artificial grammars? Transfer tests of simple association. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 396–410.

Higham, P. A. (1997). Chunks are not enough: The insufficiency of feature frequency-based explanations of artificial grammar learning. *Canadian Journal of Experimental Psychology, 51,* 126–137.

Higham, P. A., & Vokey, J. R. (2000). The controlled application of a strategy can still produce automatic effects: Reply to Redington (2000). *Journal of Experimental Psychology: General, 129,* 476–480.

Higham, P. A., Vokey, J. R., & Pritchard, J. L. (2000). Beyond dissociation logic: Evidence for controlled and automatic influences in artificial grammar learning. *Journal of Experimental Psychology: General, 129,* 457–470.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments & Computers, 16,* 96–101.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93,* 411–428.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.

Jamieson, R. K, & Mewhort, D. J. K. (2004, June). *An instance theory of structural learning*. Poster presented at the annual meeting of the Canadian Society for Brain, Behaviour, and Cognitive Sciences, St. John's, Newfoundland, Canada.

Johnstone, T., & Shanks, D. R. (1999). Two mechanisms in implicit artificial grammar learning? Comment on Meulemans and Van der Linden (1997). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 524–531.

Johnstone, T., & Shanks, D. R. (2001). Abstractionist and processing accounts of implicit learning. *Cognitive Psychology, 42,* 61–112.

Jones, M. N. (2001). *T-REX: A template-resonance excitation model of single letter classification*. Unpublished master's thesis, Queen's University, Kingston, Ontario, Canada.

Kinder, A., & Shanks, D. R. (2001). Amnesia and the declarative/nondeclarative distinction: A recurrent network model of classification, recognition, and repetition priming. *Journal of Cognitive Neuroscience, 13,* 648–669.

Kinder, A., & Shanks, D. R. (2003). Neuropsychological dissociations between priming and recognition: A single-system connectionist account. *Psychological Review, 110,* 728–744.

Knowlton, B. J., & Squire, L. R. (1992). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science, 3,* 172–179.

Knowlton, B. J., & Squire, L. R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 79–91.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22,* 169–181.

Kwantes, P., & Mewhort, D. J. K. (1999). Modeling lexical decision and word naming as a retrieval process. *Canadian Journal of Experimental Psychology, 53,* 306–315.

Mandler, G. (1967). Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: I* (pp. 327–372). Oxford, England: Academic Press.

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

Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R., & Druhan, B. (1989). The role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1083–1100.

McAndrews, M. P., & Moscovitch, M. (1985). Rule-based and exemplar-based classification in artificial grammar learning. *Memory & Cognition, 13,* 469–475.

Meulemans, T., & Van der Linden, M. (2003). Implicit learning of complex information in amnesia. *Brain & Cognition, 52,* 250–257.

Miller, G. A. (1953). What is information measurement? *American Psychologist, 8,* 3–11.

Miller, G. A. (1956a). Information and memory. In S. Coopersmith (Ed.), *Frontiers of psychological research: Readings from Scientific American* (pp. 161–164). San Francisco: Freeman.

Miller, G. A. (1956b). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Miller, G. A. (1958). Free recall of redundant strings of letters. *Journal of Experimental Psychology, 56,* 433–491.

Miller, G. A., Bruner, J. S., & Postman, L. (1954). Familiarity of letter sequences and tachistocopic identification. *Journal of General Psychology, 50,* 129–139.

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

Murdock, B. B. (1974). *Human memory: Theory and data*. Oxford, England: Erlbaum.

Neal, A., & Hesketh, B. (1997). Episodic knowledge and implicit learning. *Psychonomic Bulletin and Review, 4,* 24–37.

Perruchet, P., Bigand, E., & Benoit-Gonin, F. (1997). The emergence of explicit knowledge during the early phase of learning in sequential reaction-time tasks. *Psychological Research, 60,* 4–13.

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

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General, 119,* 264–75.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6,* 855–863.

Reber, A. S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology, 81,* 115–119.

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

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

Reber, A. S., Kassin, S. M., Lewis, S., & Cantor, G. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning & Memory, 6,* 492–502.

Schnore, M. M., & Partington, J. T. (1967). Immediate memory for visual patterns: Symmetry and amount of information. *Psychonomic Science, 8,* 421–422.

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

Shanks, D. R., & Perruchet, P. (2002). Dissociation between priming and recognition in the expression of sequential knowledge. *Psychonomic Bulletin & Review, 9,* 362–367.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17,* 367–447.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.

Stadler, M. A. (1992). Statistical structure and implicit serial learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 318–327.

Stadler, M. A. (1995). Role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21,* 674–685.

Stadler, M. A., & Neely, C. B. (1997). Effects of sequence length and structure on implicit serial learning. *Psychological Research, 60,* 14–23.

St. John, M. F., & Shanks, D. R. (1997). Implicit learning from an information processing standpoint. In D. C. Berry (Ed.), *How implicit is implicit learning?* (pp. 162–194) Oxford, England: Oxford University Press.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review, 69,* 344–354.

Tulving, E. (1985). *Elements of episodic memory*. New York: Oxford University Press.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning & Verbal Behavior, 5,* 381–391.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 359–380.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18,* 328–344.

Vokey, J. R., & Brooks, L. R. (1994). Fragmentary knowledge and the processing-specific control of structural sensitivity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20,* 1504–1510.

Vokey, J. R., & Higham, P. A. (2004). Opposition logic and neural network models in artificial grammar learning. *Consciousness and Cognition, 13,* 565–578.

Whittlesea, B. W. A., & Dorken, M. D. (1993). Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General, 122,* 227–248.

Whittlesea, B. W. A., & Wright, R. L. (1997). Implicit (and explicit) learning: Acting adaptively without knowing the consequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 181–200.

Wright, R. L., & Whittlesea, B. W. A. (1998). Implicit learning of complex structures: Active adaptation and selective processing in acquisition and application. *Memory & Cognition, 26,* 402–420.

Received April 15, 2004
Revision received July 26, 2004
Accepted July 30, 2004 ■