THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2011, 64 (2), 209–216

Psychology Press
Taylor & Francis Group

Rapid communication

Grammaticality is inferred from global similarity: A reply to Kinder (2010)

Randall K. Jamieson[1] **and D. J. K. Mewhort**[2]

[1]Department of Psychology, University of Manitoba, Winnipeg, MB, Canada
[2]Department of Psychology, Queen's University at Kingston, Kingston, ON, Canada

Jamieson and Mewhort (2009b) proposed an account of performance in the artificial-grammar judgement-of-grammaticality task based on Hintzman's (1986) model of retrieval, Minerva 2. In the account, each letter is represented by a unique vector of random elements, and each exemplar is represented by concatenating its constituent letter vectors. Although successful in simulating several experiments, Kinder (2010) showed that the model fails for three selected experiments. We track the model's failure to a constraint introduced by concatenating letter vectors to construct the exemplar representation. To fix the problem, we use a holographic representation. Holographic representation not only provides the flexibility missing with the concatenation scheme but also acknowledges variability in what subjects notice when they inspect training exemplars. Armed with holographic representations, we show that the model successfully captures the three problematic data sets. We argue for retrospective accounts, like the present one, that acknowledge subjects' skill in drawing unexpected inferences based on memory of studied items against prospective accounts that require subjects to learn statistical regularities in the training set in anticipation of an undefined classification test.

Keywords: Artificial grammar; Exemplar theory; Holographic reduced representation; Judgement of grammaticality; Global similarity.

In a judgement-of-grammaticality (JOG) task, subjects inspect strings of letters constructed using a grammar. After inspecting the strings, they classify novel strings as grammatical or ungrammatical, usually achieving about 60% correct.

Jamieson and Mewhort (2009a) explained JOG by extending Hintzman's (1986) Minerva 2 model of human memory. According to their account, subjects store each exemplar as a separate trace in memory. When a test probe is presented, each trace is activated in proportion to its similarity to the probe, and JOG is based on the similarity of the probe to an aggregate of the activated traces. The account has roots in Brooks's (1978; see also

Correspondence should be addressed to Randall K. Jamieson, Department of Psychology, University of Manitoba, Winnipeg, MB R3T 2N2, Canada. E-mail: randy_jamieson@umanitoba.ca

The research was supported by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada. We thank Michael N. Jones for comments and discussion.

http://www.psypress.com/qjep DOI:10.1080/17470218.2010.537932

Vokey & Brooks, 1992) analysis of performance in JOG.

Whereas the model captures judgement-of-grammaticality in several experiments (Jamieson, Holmes, & Mewhort, 2010; Jamieson & Mewhort, 2009a, 2010), Kinder (2010) showed that the model fails for three others. Based on the failure, Kinder argued that the Minerva 2 model is a poor account of the cognitive and memorial processes that underlie peoples' judgements-of-grammaticality.

In this reply, we sketch the model, describe why it failed, and show that it captures Kinder's (2010) examples when new representation assumptions are provided.

Jamieson and Mewhort's (2009a) model

In the model, a letter is represented by a unique vector of random elements; letter-strings are represented by concatenating letter-vectors in the same order as the letters. To illustrate, a string *ABCD* is represented by, first, generating a vector for each letter—**a**, **b**, **c**, and **d**, respectively—and, then, concatenating the vectors accordingly: *ABCD* = **a** // **b** // **c** // **d** (where // indicates concatenation). The representation scheme encodes serial-position information in the string.

Memory, **M**, is an $m \times n$ matrix. Each of the i rows in memory stores a studied exemplar. Each of the j columns represents a feature.

Storage in memory is represented by copying items to rows in **M**. Imperfect encoding is simulated by resetting a proportion of elements in **M** to zero (indicating data loss). The amount of data loss is controlled by a parameter L that specifies the probability of storing a feature in memory correctly; thus, each element in **M** has a probability $1 - L$ of reverting to zero.

When a probe is presented to memory, it is compared against all traces in parallel. Trace i's activation, a_i, is computed as,

$$a_i = \left(\frac{\sum_{j=1}^{n} p_j \times M_{ij}}{\sqrt{\sum_{j=1}^{n} p_j^2} \times \sqrt{\sum_{j=1}^{n} M_{ij}^2}} \right)^3 \qquad 1$$

where **p** is the probe, **M** is memory, i indexes the $1 \ldots m$ traces in memory, and j indexes the $1 \ldots n$ elements in the probe and traces.

The information retrieved from memory is a vector, **c**, that is the sum of the activated traces. Each of the $j = 1 \ldots n$ elements in **c** is computed as,

$$c_j = \sum_{i=1}^{m} a_i \times M_{ij} \qquad 2$$

where i indexes the $1 \ldots m$ rows in memory, and j indexes the $1 \ldots n$ columns in memory.

The model judges the grammaticality of a probe by its similarity to **c**—that is, cos{**p**,**c**}.

Why the model failed

The model captures data from several JOG experiments (Jamieson & Mewhort, 2009a, 2010; Jamieson et al., 2010) and, with a simple extension, captures performance in serial-response tasks (Jamieson & Mewhort, 2009b). However, the model fails for data from three JOG experiments. These failures reflect constraints of the concatenation-based representation scheme.

Concatenation-based representations bind letters to their serial positions. Because a trace's activation is computed across the j features of the probe and trace (see Equation 1), the model assumes that trace activation is determined by the number of letters in corresponding serial positions. Because of the serial-position constraint, a probe fails to activate traces that share its letters, bigrams, and trigrams, unless those units appear at the identical serial positions. Thus, *ABCD* will not activate traces *CDAB*, *BCDA*, and *DBAC*. This aspect of the model contradicts data (Kinder, 2010). A more appropriate scheme would capture what people notice in the training exemplars and acknowledge overlap of units independent of their serial positions.

Holographic representation

Here, we propose a representation scheme based on the mathematics of holography. We rely on

Gabor's (1968, 1969; see also Longuet-Higgins, 1968; Poggio, 1973) insights that a representation based on vector convolution mimics a hologram and can be used to encode serial information (see also Metcalfe-Eich, 1982; Murdock, 1995). Holographic representation solves other long-standing puzzles. A holographic representation of the word *bank*, for example, can hold simultaneously the several meanings of that word (Jones & Mewhort, 2007).

Our method for holographic representation relies on circular convolution. Circular convolution is an operation that associates two vectors $\mathbf{x}$ and $\mathbf{y}$ by collapsing their outer-product matrix to form a new vector, $\mathbf{z}$:

$$z_i = \sum_{j=0}^{n-1} x_{j \bmod n} \times y_{(i-j) \bmod n} \quad \{\text{for } i = 0 \text{ to } n-1\} \tag{3}$$

Figure 1 illustrates the computation of $\mathbf{z}$ from two vectors, $\mathbf{x}$ and $\mathbf{y}$, both of dimensionality 5. The operation is commutative and associative, and distributes over addition.

Noncommutative circular convolution is used to model asymmetric serial-order information; such as the left-to-right directional associations amongst letters that subjects develop in the JOG task. Noncommutative circular convolution is accomplished by scrambling indices in a letter-vector's representation differently depending on whether it is the predecessor or successor in a bigram (see Jones & Mewhort, 2007; Plate, 1995). Plate's noncommutative circular convolution is neither commutative nor associative, but it distributes over addition and preserves similarity. In the remainder of this paper, we denote noncommutative circular convolution by an asterisk (e.g., $\mathbf{z} = \mathbf{x} * \mathbf{y}$). For brevity, we use the term *convolution* in place of noncommutative circular convolution.

To illustrate how one constructs the holographic representation of a letter-string, consider the item *<ABCD>* (following Kinder & Lotz, 2009, **<** indicates the start, and **>** indicates the end of the string). To capture first-order structure of *<ABCD>*, we sum all individual letter-vectors: $\mathbf{<} + \mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} + \mathbf{>}$. To capture second-order structure, we sum all convolved bigrams: $\mathbf{<}*\mathbf{a} + \mathbf{a}*\mathbf{b} + \mathbf{b}*\mathbf{c} + \mathbf{c}*\mathbf{d} + \mathbf{d}*\mathbf{>}$. To capture its

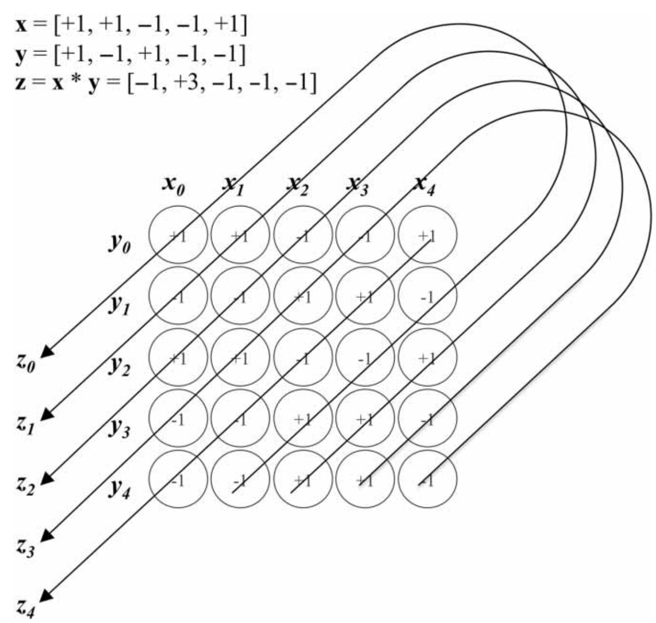

Figure 1. *The figure shows two vectors, $\mathbf{x}$ and $\mathbf{y}$ (both of dimensionality $n = 5$). The outer product of $\mathbf{x}$ and $\mathbf{y}$ is a $n \times n$ matrix. The arrows show how the elements of the outer-product matrix are summed during circular convolution to produce a summary vector $\mathbf{z}$.*

third-order structure, we sum all convolved trigrams: $(\mathbf{<}^*\mathbf{a})^*\mathbf{b} + (\mathbf{a}^*\mathbf{b})^*\mathbf{c} + (\mathbf{b}^*\mathbf{c})^*\mathbf{d} + (\mathbf{c}^*\mathbf{d})^*\mathbf{>}$. To represent $<ABCD>$ as a composite of letters, bigrams, and trigrams, we sum the units from all three levels. One could extend the operation to include larger chunks (e.g., four-grams, five-grams, and so on).

Holographic representations encode serial-order information independently from serial-position information. Because serial-order information is decoupled from serial-position information, a probe can retrieve memory traces of strings that share its letters, bigrams, and trigrams, even when the strings appear at different serial positions. This is the property of holographic representation that finesses the serial-position constraint associated with concatenation-based models.

Of course, people do not encode all information in a string, and it is unlikely that all subjects notice the same information (Wright & Whittlesea, 1998). To acknowledge these facts, we represent an exemplar as a random sample of six units (i.e., single letters, bigrams, and trigrams). Thus, in our simulations, a string $<ABCD>$ might be represented as $\mathbf{a} + \mathbf{b} + \mathbf{d} + \mathbf{<}^*\mathbf{a} + \mathbf{b}^*\mathbf{c} + (\mathbf{b}^*\mathbf{c})^*\mathbf{d}$ at one encoding and $\mathbf{b} + \mathbf{b} + \mathbf{<}^*\mathbf{a} + (\mathbf{b}^*\mathbf{c})^*\mathbf{d} + (\mathbf{<}^*\mathbf{a})^*\mathbf{b} + (\mathbf{c}^*\mathbf{d})^*\mathbf{>}$ at another.

Simulations

We applied Jamieson and Mewhort's (2009b) retrieval model against the three judgement-of-grammaticality experiments that Kinder (2010) identified as problematic. For comparison, we report simulations using both the concatenation-based representations and the holographic representations.

All of the simulations that follow were conducted using the same procedure. First, we generated representations for the relevant training and test strings. Second, we stored the training strings to memory ($L = 1$). Third, we recorded the echo intensity of each test string. We report mean echo intensity for classes of test strings; however, the model provides an echo intensity for each string.

Simulation 1: Knowlton and Squire (1996)

Knowlton and Squire (1996) reported an experiment in which subjects studied grammatical training strings and then judged the grammaticality of novel test strings. The test set included four classes of items defined factorially by two levels of grammatical status (grammatical and ungrammatical) and two levels of similarity (high and low). Similarity was defined as associative chunk strength (ACS): High-ACS items shared more bigrams and trigrams with training items than did low-ACS items.

Knowlton and Squire's (1996) results are presented in the top panel of Figure 2. As shown, subjects preferred grammatical to ungrammatical test strings and preferred high-ACS to low-ACS ungrammatical items. The selective influence of ACS on ungrammatical strings issues a constraint on models of JOG.

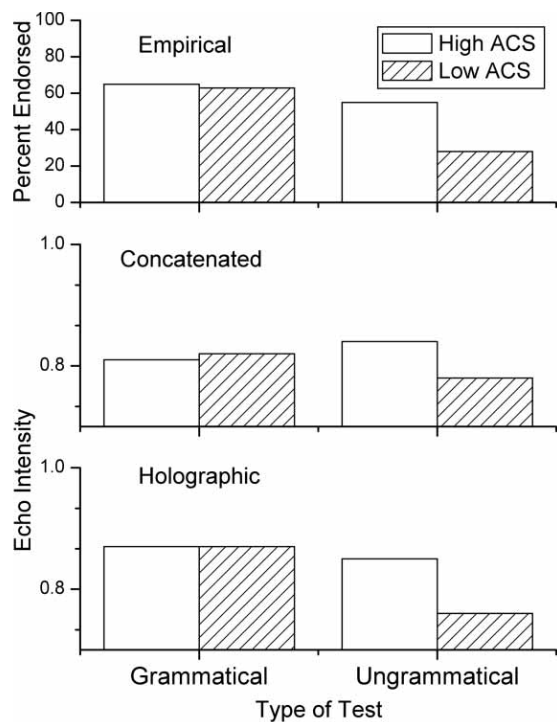

Figure 2. *Grammaticality judgement as a function of both grammatical status and associative chunk strength (ACS). The top panel shows Knowlton and Squire's (1996, Experiment 1) results. The centre panel shows results for the concatenation-based model. The bottom panel shows results for the convolution-based model. For both simulations, $L = 1$, where L is a parameter that specifies the probability of storing a feature in memory correctly.*

We conducted 100 simulations of Knowlton and Squire's (1996) experiment using concatenation-based representations and 100 simulations using holographic representations. The middle and bottom panels of Figure 2 show mean echo intensity for the four classes of items from the two sets of simulations. As shown, the two models generate qualitatively different predictions. The holographic model predicted Knowlton and Squire's data whereas the concatenation-based model did not. Most notable, the concatenation-based model failed to anticipate subjects' preference for grammatical over ungrammatical test items.

Based on our simulations, we argue that holographic representation is the more accurate method for capturing how people represent training and test strings in memory. Critical to our conclusion that JOG reflects exemplar-based similarity, we did not change our model's account of storage and retrieval from memory; only the representation scheme differed.

Simulation 2: Kinder (2000)

In the first simulation, we showed that the holographic representation scheme solves the problems associated with the concatenation-based scheme. The second simulation generalizes that argument to data from Kinder (2000; see also Kinder & Lotz, 2009).

Kinder manipulated low-order serial dependencies, high-order serial dependencies, and positional dependencies of letters in training and test strings across four types of test item. The items were designed, in psychophysical tradition, to uncover properties of the stimuli to which subjects are sensitive. Of the four types, Types 1 and 2 were ungrammatical; Types 3 and 4 were grammatical. Type 1 items displaced letters from their serial positions in the training strings; Type 2 items did not. Type 3 items shared small chunks with the training set (i.e., bigrams and trigrams); Type 4 items shared large chunks with the training set (i.e., quartets or greater). An advantage for Type 2 over Type 1 tests exposes sensitivity to positional rules for ungrammatical items. An advantage for Type 4 over Type 3 tests exposes sensitivity to large- over small-chunk similarity, independent of grammatical status.

Figure 3 (top panel) shows subjects' grammatical endorsement for the four types of test (redrawn from Kinder, 2000, Fig. 3, p. 99). As shown, subjects strongly favoured grammatical over ungrammatical test strings (i.e., Types 3 and 4 items over Types 1 and 2 items), reliably, albeit modestly, favoured ungrammatical test strings that had letters in serial positions that were consistent with the grammar's positional rules (i.e., Type 2 over Type 1 items), and (c) unreliably, but measurably, favoured grammatical tests that shared large chunks with specific training items (i.e., Type 4 over Type 3 items). A competent model should be capable of reproducing the data.

We conducted 100 simulations of Kinder's (2000) experiment using concatenation-based representations and 100 simulations using

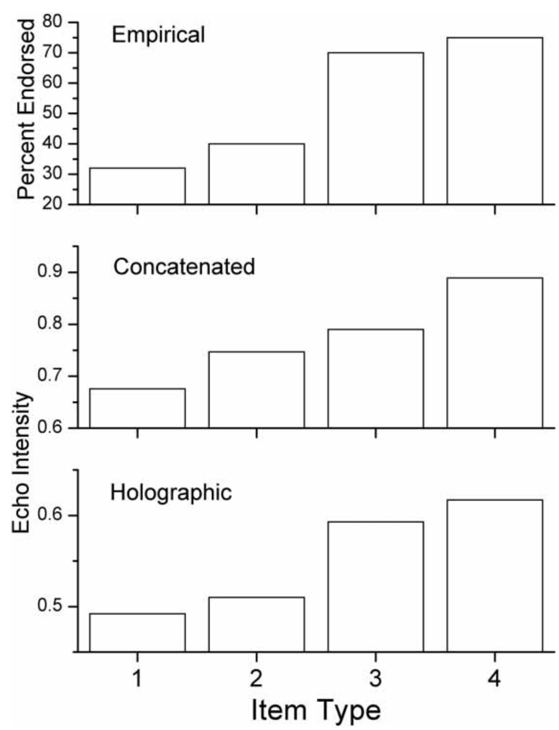

Figure 3. *Grammaticality judgement as a function of item type. The top panel shows Kinder's (2000) results. The centre panel shows results for the concatenation-based model. The bottom panel shows results for the convolution-based model. For both simulations, $L = 1$, where L is a parameter that specifies the probability of storing a feature in memory correctly.*

holographic representations. The middle panel of Figure 3 shows mean echo intensity for the four classes of items using the concatenation-based representations; the bottom panel of Figure 3 shows the mean echo intensity for the four classes of test items using the holographic representations. As shown, the holographic model predicted Kinder's data whereas the concatenation-based model did not. Like Kinder's subjects, the holographic model exhibits a strong preference for grammatical over ungrammatical strings (i.e., Types 3 and 4 over Types 1 and 2), a weak preference for ungrammatical strings that conformed to the grammar's positional rules (i.e., Type 2 over Type 1), and a weak preference for grammatical test strings that included a near neighbour in the training set (i.e., Type 4 over Type 3 strings). Unlike her subjects, the concatenation-based model exhibited a strong sensitivity to violations of the grammar's positional dependencies (i.e., Type 2 over Type 1 items) and a strong endorsement of grammatical strings that shared large chunks with specific training items (i.e., Type 4 over Type 3 items). The simulation reaffirms our conclusions from Simulation 1.

Simulation 3: Kinder and Shanks (2001)

In the first two simulations, we showed that the holographic representation scheme gives a better fit to data than does the concatenation-based scheme. Simulation Study 3 examines the same problem using data from Kinder and Shanks (2001).

Kinder and Shanks (2001) included unstudied grammatical items, studied grammatical items, and unstudied ungrammatical items. Figure 4 (top panel) shows subjects' grammatical endorsement for the three classes of items (drawn from Kinder & Shanks, 2001, Table 1, p. 653). As shown, subjects preferred grammatical over ungrammatical strings but did not distinguish studied from the unstudied grammatical test items. The result suggests that a competent model of the artificial grammar task should not prefer studied to unstudied strings; however, see Jamieson and Mewhort (2010) for data that suggest the opposite (we return to this point in the Discussion section).

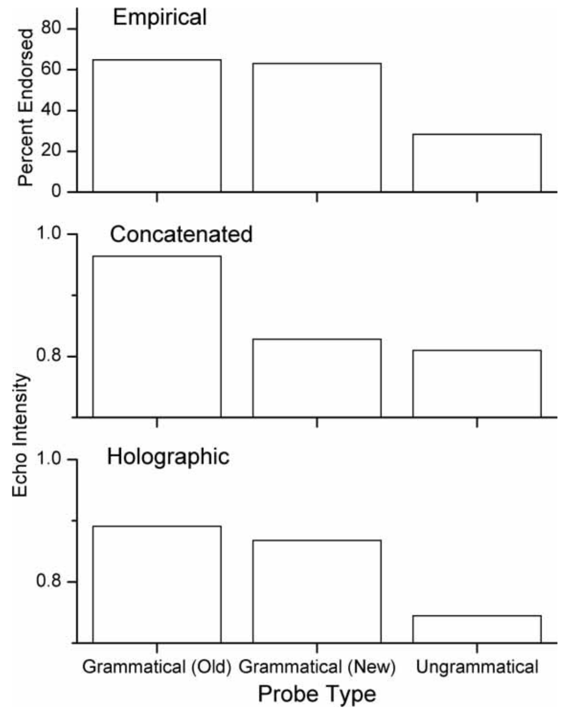

Figure 4. *Grammaticality judgement as a function of both grammatical and study status. The top panel shows Kinder and Shanks' (2001, Experiment 1) results. The centre panel shows results for the concatenation-based model. The bottom panel shows results for the convolution-based model. For both simulations, L = 1, where L is a parameter that specifies the probability of storing a feature in memory correctly.*

We conducted simulations of Kinder and Shanks's (2001) experiment using concatenation-based and holographic representations. The middle panel of Figure 4 shows results for the concatenation-based model; the lower panel shows results with the holographic model. Whereas the concatenation-based model incorrectly predicts that subjects should prefer studied to unstudied items, the holographic model correctly predicts that subjects prefer grammatical over ungrammatical items. Once again, the holographic model does a better job of predicting subjects' judgements-of-grammaticality.

Discussion

Kinder (2010) argued that Jamieson and Mewhort's (2009a) retrieval model is wrong. We have traced the problem to our previous and naïve representation scheme. The concatenation scheme encoded a string of letters as if the letters were assigned to serial positions. The holographic scheme assumes that participants encode a string as a collection of its subunits, retaining details of serial-order information in the training items, but very little about serial-position information—a position consistent with data (e.g., Kinder, 2010). Armed with holographic representations, the model accommodates the results of all three experiments that Kinder identified as problematic for it. The holographic model also captures data from previous simulations (Jamieson et al., 2010; Jamieson & Mewhort, 2009a, 2010); because of space limitations, we do not report those simulations here. We conclude that holographic representation is a more accurate method for capturing what people notice in training and test strings.

By changing the representations, one might object that we have created a new model. In our view, representation assumptions are important but not definitional. The model's core assumptions still apply: Subjects store aspects of each training exemplar in memory and infer a test probe's grammaticality from its global similarity to the training set. The difference is that the information noticed is stored in a distributed fashion within a holographic structure. We see no need to assume that the subject extracts knowledge of the grammar or tabulates the frequency distributions of chunks (or other units) across the studied exemplars. Admittedly, however, our evidence does not rule out such learning.

Our retrieval-based explanation for JOG contrasts to learning explanations. Learning accounts assume that the subject prepares for the test. However, because the subjects are informed of the test only after they have inspected the training set, it is not clear how the subject would intuit the kind of test they will be given. By contrast, our memory account assumes that subjects notice aspects of the exemplars and, at test, use their memory of the training strings. If the probe reminds them of studied exemplars, they judge it to be grammatical. In Redington and Chater's (2002) terms, learning accounts are *prospective* and *eager* whereas retrieval accounts, like ours, are *retrospective* and *lazy*. Stepping aside from model fitting, we prefer a retrieval-based explanation because it more closely matches the demand/constraints of the experimental procedure.

In Kinder and Shanks's (2001) experiment, subjects showed only a modest preference for studied over unstudied test strings. We have observed results both consistent and inconsistent with that position (Jamieson & Mewhort, 2010). The distinction points to a problem in asserting that subjects respond to one form of structure over another and, thus, to any theory of JOG that asserts that subjects learn a specific form of information. Wright and Whittlesea (1998) demonstrated the problem nicely using digit strings. When strings were arrayed horizontally, judgements were consistent with the numerical distance between strings; when arrayed vertically, judgements were consistent with a match of digits in corresponding serial positions. Thus, even mundane stimulus manipulations change the encoding strategies that subjects adopt and, consequently, their judgements at test based on memory of the exemplars. For example, if string length exceeds memory span, subjects might encode strings as subunits rather than as whole strings (Jamieson & Mewhort, 2005; Miller, 1956). We acknowledge that encoding flexibility makes a mechanistic account of JOG difficult; however, we also see it to be a behavioural fact that needs to be understood (Jamieson & Mewhort, 2005). The holographic model accepts variability at encoding and provides a scheme to represent it. We have used a random sample of subunits here; however, we can generate predictions for any systematic encoding strategy. In our view, models of JOG must accommodate flexibility in subjects' use of information, rather than explain how subjects use a particular kind of information over all others.

Original manuscript received 18 August 2010
Accepted revision received 12 October 2010

REFERENCES

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: Lawrence Erlbaum Associates.

Gabor, D. (1968). Improved holographic model of temporal recall. *Nature*, *217,* 1288–1289.

Gabor, D. (1969). Associative holographic memories. *IBM Journal of Research and Development*, *13,* 156–159.

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

Jamieson, R. K., Holmes, S., & Mewhort, D. J. K. (2010). Global similarity predicts dissociation of classification and recognition: Evidence questioning the implicit/explicit learning distinction in amnesia. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *36*, 1529−1535.

Jamieson, R. K., & Mewhort, D. J. K. (2005). The influence of grammatical, local, and organizational redundancy on implicit learning: An analysis using information theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31,* 9–23.

Jamieson, R. K., & Mewhort, D. J. K. (2009a). Applying an exemplar model to the artificial-grammar task: Inferring grammaticality from similarity. *Quarterly Journal of Experimental Psychology*, *62,* 550–575.

Jamieson, R. K., & Mewhort, D. J. K. (2009b). Applying an exemplar model to the serial reaction-time task: Anticipating from experience. *Quarterly Journal of Experimental Psychology*, *62,* 1757–1783.

Jamieson, R. K., & Mewhort, D. J. K. (2010). Applying an exemplar model to the artificial-grammar task: String-completion and performance on individual items. *Quarterly Journal of Experimental Psychology*, *63,* 1014–1039.

Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, *114,* 1–37.

Kinder, A. (2000). The knowledge acquired during artificial grammar learning: Testing the predictions of two connectionist models. *Psychological Research*, *63,* 95–105.

Kinder, A. (2010). Is grammaticality inferred from global similarity? Comment on Jamieson & Mewhort (2009). *Quarterly Journal of Experimental Psychology*, *63,* 1049–1056.

Kinder, A., & Lotz, A. (2009). Connectionist models of artificial grammar learning: What type of knowledge is acquired? *Psychological Research*, *73,* 659–673.

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.

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, and Cognition*, *22,* 169–181.

Longuet-Higgins, H. C. (1968). Holographic model of temporal recall. *Nature*, *217,* 104.

Metcalfe-Eich, J. (1982). A composite holographic associative recall model. *Psychological Review*, *89,* 627–661.

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

Murdock, B. B. (1995). Developing TODAM: Three models for serial-order information. *Memory & Cognition*, *23,* 631–645.

Plate, T. A. (1995). Holographic reduced representations. *IEEE Transactions on Neural Networks*, *6,* 623–641.

Poggio, T. (1973). On holographic models of memory. *Kybernetik*, *12,* 237–238.

Redington, M., & Chater, N. (2002). Knowledge representation and transfer in artificial grammar learning (AGL). In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness: An empirical, philosophical and computational consensus in the making* (pp. 121–143). Hove, UK: Psychology Press.

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

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.