

# Sources of confidence judgments in implicit cognition

RICHARD J. TUNNEY

*University of Nottingham, Nottingham, England*

Subjective reports of confidence are frequently used as a measure of awareness in a variety of fields, including artificial grammar learning. However, little is known about what information is used to make confidence judgments and whether there are any possible sources of information used to discriminate between items that are unrelated to confidence. The data reported here replicate an earlier experiment by Vokey and Brooks (1992) and show that grammaticality decisions are based on both the grammatical status of items and their similarity to study exemplars. The key finding is that confidence ratings made on a continuous scale (50%–100%) are closely related to grammaticality but are unrelated to all of the measures of similarity that were tested. By contrast, confidence ratings made on a binary scale (high vs. low) are related to both grammaticality and similarity. The data confirm an earlier finding (Tunney & Shanks, 2003) that binary confidence ratings are more sensitive to low levels of awareness than continuous ratings are and suggest that participants are conscious of all the information acquired in artificial grammar learning.

For many years, subjective reports of confidence have been used to distinguish between conscious and unconscious cognitive processes in areas as diverse as discrimination (Peirce & Jastrow, 1884), perception (Cheesman & Merikle, 1984; Kunimoto, Miller, & Pashler, 2001), and memory (Yonelinas, 1997). However, much of the recent interest in unconscious cognition has been in learning, and it has been studied in some depth, using the artificial grammar learning paradigm (Reber, 1967). Despite this, little is known about what information is used to make confidence judgments when participants are aware and whether there are any possible sources that they use to make a discrimination that they might be unaware of.

In a typical artificial grammar learning experiment, participants study strings of letters that, unknown to them, have been generated by an artificial grammar. In a subsequent test, the participants are asked to discriminate between new grammatical and ungrammatical items. Reber (1967) originally argued that knowledge of the grammar used to classify grammatical and ungrammatical sequences was implicit, because participants could not articulate the basis of their decision. Verbal reports of awareness are frequently criticized because they are prone to be influenced by response bias (e.g., Eriksen, 1960; Shanks & St. John, 1994); participants may claim to be unaware, when in fact they possess a degree of awareness, albeit small, of the knowledge they are using. A participant may set a conservative response criterion and choose to report only knowledge held with high confidence. Knowl-

edge with low confidence may, therefore, not be reported even if, in fact, conscious. An alternative to verbal report is to ask participants to rate how confident they are that their judgments are correct or, alternately, to estimate their overall level of accuracy.

Several techniques for assessing levels of awareness on the basis of subjective confidence have been used (see Dienes & Berry, 1997). The *guessing criterion* defines knowledge as unconscious when participants believe they are guessing. For example, in each of Dienes, Altmann, Kwan, and Goode's (1995) five experiments, participants reported a proportion of their responses to be guesses (i.e., their confidence was 50% on a scale of 50%–100%). Dienes et al. found that, reliably, more guesses were correct than were incorrect or, to put it another way, that classification accuracy based only on guesses was greater than chance. Dienes et al. concluded that artificial grammar learning involves some unconscious knowledge, according to a subjective criterion of awareness based on guessing.

The *zero-correlation criterion* defines awareness in terms of the relationship between confidence and accuracy. Chan (1992) suggested that if participants are aware of the knowledge they use to classify items, they should be more confident when they make correct decisions than when they make incorrect ones. Thus, confidence should predict accuracy. Conversely, if discriminations are made without awareness, participants should be just as confident when they make incorrect decisions as when they make correct ones. Data from experiments using this criterion suggest that learning can be implicit under some circumstances. For instance, Dienes and Altmann (1997) reported that when participants classified strings in the same vocabulary as that used during learning, there was a small but reliable relationship between con-

---

Correspondence concerning this article should be addressed to R. J. Tunney, School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, England (e-mail: rjt@psychology.nottingham.ac.uk).

fidence and accuracy, suggesting a degree of explicit knowledge. In contrast, when the vocabulary of the test strings was changed, the relationship between confidence and accuracy disappeared. Dienes and Altmann concluded that the knowledge used to classify strings in a novel vocabulary was implicit. Recently, Tunney and Shanks (2003) refined this technique by asking participants to report their confidence, using a binary scale (high vs. low). This method cannot be used to assess the accuracy of guesses but can be used in precisely the same way as the zero-correlation criterion. Tunney and Shanks reported that when the zero-correlation criterion was used to compare the two scales, the binary one was more sensitive to low levels of awareness than the continuous scale was, in the sense that binary confidence ratings indicated explicit knowledge, whereas continuous confidence ratings did not.

These criteria of awareness assume that if participants are aware of the knowledge they use to classify items, they should be more confident in correct than in incorrect decisions. It follows that the information used to make confidence ratings should be the same as that used to classify items. A number of studies suggest that participants classify test items on the basis of rules and on the basis of similarity to study items.

For instance, in an elegant experiment, Vokey and Brooks (1992) found evidence of multiple sources of grammaticality decisions. The stimuli used in this experiment were constructed in such a way that the similarity of the test items was orthogonal to their grammatical status. Vokey and Brooks used a nearest neighbor measure of similarity: Half the test items were similar to the training items, and differed from a paired study item by a single letter. The other half was dissimilar and differed by two letters or more. Half of each list conformed to the grammatical rules, and half did not. Although the participants were instructed to classify items according to grammatical status, Vokey and Brooks found that both similarity and grammaticality had an additive effect on endorsement rates (see also McAndrews & Moscovitch, 1985). This raises the possibility that participants learn some grammatical rules and remember some specific items.

Higham (1997b) showed that memory- and rule-based processes could be dissociated by dividing attention at test. Dividing attention reduced the contribution of similarity, but not the contribution of grammaticality, to classification performance. Higham (1997b) suggested that dividing attention impaired the use of conscious recollection but had no effect on implicit abstraction processes. In a similar vein, Meulemans and Van der Linden (1997) showed that memory for study items and, consequently the contribution of similarity to test item classification could be increased by reducing the number of items to be memorized. Conversely, increasing the number of items to be memorized enhances the contribution of grammaticality to classification.

Finally, data from amnesic patients are frequently cited as evidence that memory and rule abstraction are dissocia-

ble processes in artificial grammar learning. For example, Knowlton, Ramus, and Squire (1992) compared the ability of amnesic patients and age-matched controls to classify items according to grammatical status and recognize the study exemplars. They found that although recognition was impaired in the amnesic group, classification performance was similar between the two groups (see Kinder & Shanks, 2001, for an alternative interpretation).

These data suggest that classification is based on explicit memory of study exemplars and implicit knowledge of grammatical rules. Little is known, however, of what information participants use to estimate their accuracy. More specifically, what are the sources of the confidence judgments used to measure awareness? The issue is important because, if discrimination accuracy is calculated over information that is imperfectly related to the information participants actually use to make a discrimination, it is likely to be even less perfectly related to the index of awareness. Indices of awareness based on confidence ratings are usually derived from the difference in confidence for correct and incorrect decisions, where a correct decision is defined as one based on the grammaticality of the test items. Such indices tend to indicate that participants have little or no awareness in the accuracy of their decisions (Dienes & Berry, 1997). However, there is clear evidence that similarity plays a major role in decisions about category membership (Perruchet & Pacteau, 1990; Shanks & St. John, 1994) and that these decisions may be predicated on explicit knowledge. If participants classify test items on the basis of their similarity to study items and rate their confidence accordingly, attempts to relate these variables to accuracy calculated over the items' grammatical status would return an imperfect measure of awareness. This would inevitably lead to the erroneous conclusion that artificial grammar learning is largely an implicit process. The experiment that follows is, in essence, a replication of the experiment reported by Vokey and Brooks (1992), in which the similarity of test items was orthogonal to their grammaticality. The novel aspect of this replication was that the participants rated their confidence in each decision on two scales (continuous and binary), permitting subjective measures of awareness to be calculated for decisions based on both grammaticality and similarity. The question is, are participants aware of all the information that they use to classify test items?

## METHOD

The experiment proceeded in the same way as a typical AGL experiment, with the exception that, unknown to the participants, the test sequences could be split into four orthogonal partitions: grammatical-similar, grammatical-dissimilar, ungrammatical-similar, and ungrammatical-dissimilar. Grammatical items could be generated by the same grammar that generated the training exemplars; ungrammatical items could not. Similarity was defined using a nearest neighbor criterion: each similar item differed from training item by a single letter, and dissimilar items differed by two or more letters. The novel aspect of this replication was that the participants rated their confidence in each decision, permitting sub-

jective measures of awareness to be calculated for decisions based on both grammaticality and similarity.

### Participants

Forty-two undergraduate students took part in this experiment as part of a course requirement. Fourteen were male, and 28 were female (mean age = 19.42 years,  $SD = 2.37$ ).

### Stimuli

There were two lists of training items and four lists of test items, shown in the Appendix. These were the same items as those described by Vokey and Brooks (1992). To control for any possible biases in the items, the participants were presented with either List 1 or List 2 training items. Each list consisted of 16 items and was presented twice in a different random order for each participant. The same List 1 and List 2 items were used for each participant. At test, all the participants were asked to classify 64 items drawn from both List 1 and List 2. For the participants who had studied List 1, the List 1 test items were similar (differed by a single letter), and the List 2 items were dissimilar (differed by two or more letters). Conversely, for the participants who had studied List 2, the List 2 test items were similar, and the List 1 items were dissimilar. For all the groups, half the test items were grammatical, and half were ungrammatical.

### Procedure

The participants were asked to memorize each training item. Training items appeared on a computer screen for 5 sec. After this, the item disappeared, and the participants were asked to use the computer keyboard to type the item that would appear in the same location on the screen. If the item was typed incorrectly, it would reappear for a further 5 sec; otherwise, the next item would appear. This procedure differed slightly from that used by Vokey and Brooks (1992, Experiment 1), who asked participants to memorize and then free-recall training items in blocks of four.

At test, the participants were informed that all the items in the first part of the experiment were constructed using a simple set of rules and that their task was to decide which of a new set of items were constructed using the same rules. The participants were informed that half of the items were constructed using the same rules. The grammaticality decision was made on a 4-point scale that enabled confidence to be reported on a binary scale concurrently: the participants clicked one of four buttons on screen marked *yes conforms to rules—more confident*, *yes conforms to rules—less confident*, *no does not conform to rules—more confident*, and *no does not conform to rules—less confident*. After this, the participants were asked to rate their decision on a continuous scale from 50% to 100%, where 50% corresponded to a *complete guess* and 100% corresponded to *absolute certainty*. No feedback concerning discrimination accuracy was provided.

## RESULTS

The proportion of correct grammaticality judgments was .58 ( $SE = .01$ ). This was reliably greater than an expected chance value of .5 [ $t(41) = 8.55, t < .01$ ]. Table 1 shows the mean proportion of items endorsed as grammatical as a function of grammaticality and similarity. These data were entered into a  $2 \times 2$  (item vs. similarity) repeated measures analysis of variance (ANOVA). There was an effect of item, indicating that the participants endorsed grammatical items more often than ungrammatical items [ $F(1,41) = 73.87, MS_e = 0.02, p < .01$ ], an effect of similarity, indicating that the participants endorsed similar items more often than dissimilar items [ $F(1,41) = 32.33, MS_e = 0.01, p < .01$ ], but no interaction [ $F(1,41) = 2.12, MS_e = 0.01, p > .05$ ]. These data closely replicate those reported by Vokey and Brooks (1992).

On what basis did the participants make their confidence ratings? Table 1 also shows the mean confidence reported on the continuous scale and the mean number of high-confidence responses reported on the binary scale, as a function of response (whether the participants endorsed items as grammatical or rejected items as ungrammatical), item (whether the item was grammatical or ungrammatical), and similarity (whether the item was similar or dissimilar to training items).

### Continuous Confidence

The mean confidence ratings were entered into a  $2 \times 2 \times 2$  (response, item, and similarity) repeated measures ANOVA. This revealed an effect of response, indicating that the participants were more confident when they endorsed items as grammatical than when they rejected items as ungrammatical [ $F(1,41) = 13.01, MS_e = 84.01, p < .01$ ]. There was an effect of item, indicating that the participants were more confident when classifying grammatical than when classifying ungrammatical items [ $F(1,41) = 15.30, MS_e = 14.29, p < .01$ ], and an interaction between the two [ $F(1,41) = 13.58, MS_e = 23.43, p < .01$ ]. The interaction was due to an effect of response for grammatical items, indicating that the participants were more confident when they correctly endorsed grammatical items than when they were incorrectly rejected [ $F(1,83) =$

**Table 1**  
Mean Proportions of Endorsements, Mean Confidence Given on the Continuous Confidence Scale, and Mean Numbers of High-Confidence Responses Given on the Binary Confidence Scale, as a Function of Grammaticality and Similarity

	Grammatical				Ungrammatical			
	Similar		Dissimilar		Similar		Dissimilar	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Discrimination								
Endorsements	.66	.02	.59	.02	.52	.03	.41	.02
Continuous confidence								
Endorsements	74.50	1.37	75.59	1.23	71.48	1.29	71.48	1.13
Rejections	69.36	1.65	69.59	1.55	69.07	1.51	70.54	1.44
Binary confidence								
Endorsements	5.26	0.44	4.88	0.36	3.38	0.31	2.45	0.30
Rejections	1.91	0.26	2.24	0.30	2.98	0.36	3.62	0.40

29.45,  $MS_e = 44.19$ ,  $p < .01$ ]. There was no effect of response for ungrammatical items [ $F(1,83) = 3.61$ ,  $MS_e = 32.63$ ,  $p > .05$ ]. This pattern is shown in Figure 1A.

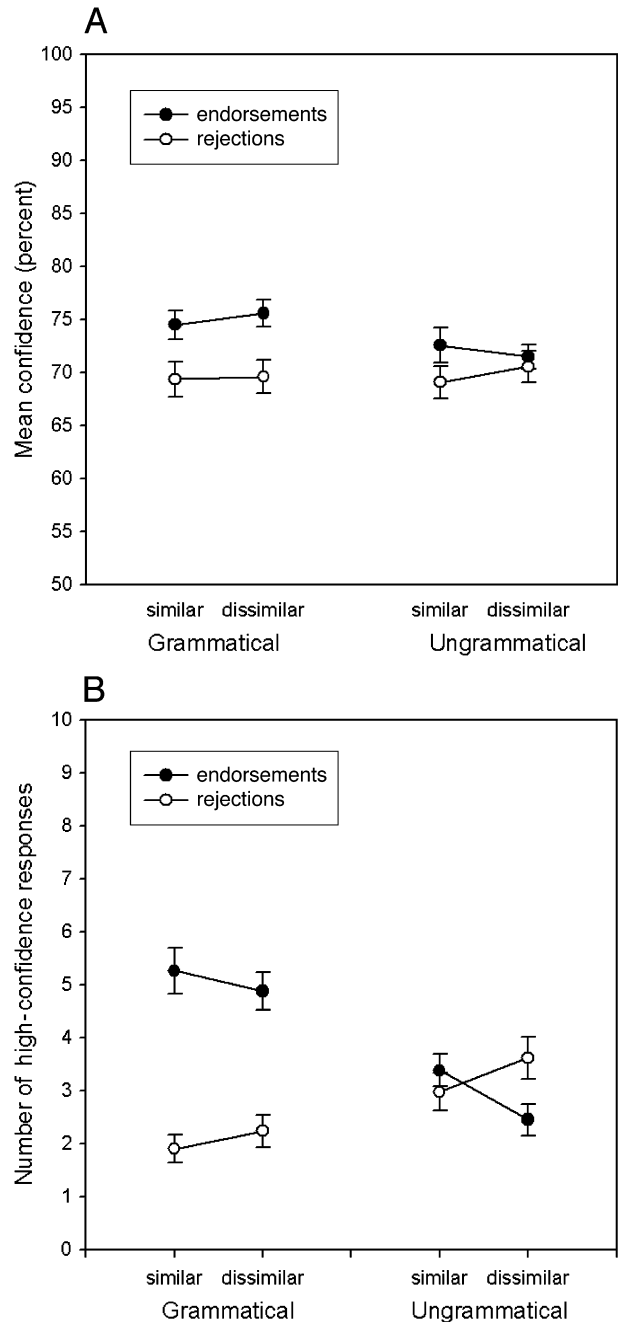
There was no effect of similarity [ $F(1,41) = 1.62$ ,  $MS_e = 25.30$ ,  $p > .05$ ] and no interaction with either response or item (both  $F$ s  $< 1.0$ ), indicating that continuous confidence ratings were uninfluenced by this measure of item similarity.

### Binary Confidence

Previous research has shown that confidence ratings made on a binary scale are more sensitive to low levels of awareness than ratings made on a continuous scale are (Kunimoto et al., 2001; Tunney & Shanks, 2003). If so, then binary confidence ratings might show an effect of similarity where continuous ratings do not. For the most part, the binary confidence ratings show the same pattern as the continuous confidence ratings. There was an effect of response [ $F(1,41) = 12.61$ ,  $MS_e = 11.43$ ,  $p < .01$ ], an effect of item [ $F(1,41) = 12.19$ ,  $MS_e = 1.49$ ], and an interaction between the two [ $F(1,41) = 65.20$ ,  $MS_e = 3.68$ ,  $p < .01$ ]. As with the continuous ratings, this interaction was due to an effect of response for grammatical items, because the participants were more confident when they correctly endorsed grammatical items than when they incorrectly endorsed ungrammatical items [ $F(1,83) = 77.28$ ,  $MS_e = 4.89$ ,  $p < .01$ ]. There was no effect of response for ungrammatical items [ $F(1,83) = 1.23$ ,  $MS_e = 4.95$ ,  $p > .05$ ]. There was no effect of similarity ( $F < 1.0$ ), but in contrast to the continuous confidence ratings, there was an interaction between similarity and response [ $F(1,41) = 13.62$ ,  $MS_e = 2.01$ ,  $p < .01$ ]. There was no interaction between similarity and item ( $F < 1.0$ ). The three-way interaction failed to reach significance [ $F(1,41) = 1.89$ ,  $MS_e = 2.04$ ,  $p > .05$ ]. An examination of the interaction between similarity and response revealed that the number of high-confidence responses increased with increases in similarity for endorsements [ $F(1,83) = 8.61$ ,  $MS_e = 2.09$ ,  $p < .01$ ] and decreased with increases in similarity for rejections [ $F(1,83) = 5.52$ ,  $MS_e = 1.81$ ,  $p < .05$ ]. This pattern is shown in Figure 1B.

### Alternative Measures of Similarity

Although continuous confidence ratings were related to grammaticality, they were unrelated to item similarity. By contrast, binary confidence did show an interaction between similarity and response. However, the measure of similarity used by Vokey and Brooks (1992), and replicated here, has been criticized because it may not correspond to empirically derived indices of similarity (Knowlton & Squire, 1994; Perruchet, 1994; Pothos & Bailey, 2000). Indeed, a good deal of research suggests that participants categorize test items according to their similarity computed over letter chunks (bigrams and trigrams), rather than whole items. It is possible that continuous confidence ratings reflect a different index of similarity than binary ratings do. To determine the extent to which chunk strength influences participants' confidence rat-



**Figure 1.** (A) Mean continuous confidence ratings by item grammaticality, response, and similarity. (B) Mean number of high-confidence responses by item grammaticality, response, and similarity.

ings, three measures of chunk strength were calculated that have previously been shown to influence categorization (Johnstone & Shanks, 1999, 2001): *bigram chunk strength* (BCS), *anchor position chunk strength* (ACS), using the method described by Higham (1997a) and Knowlton and Squire (1994), and *first and last letter strength* (FLS), using a method similar to that described

by Tunney and Altmann (1999). The first, BCS, is a measure of how similar each test string is to the training list on the basis of the number of times the bigrams in each string occurred in the training list. The second, ACS, is calculated on the basis of the initial and terminal bigrams. The third, FLS, is simply the frequency of occurrence of letters in the first and last positions of each sequence. Using the same method described by Johnstone and Shanks (1999) separate regressions were calculated for each participant, with indices of similarity as predictors of confidence ratings. The mean slope and related *t* tests are shown in Table 2. The number of high-confidence responses was reliably predicted by each measure of similarity. The number of high-confidence responses was positively related to similarity (except ACS) for endorsements, and negatively related for rejections. Fewer measures of similarity were reliable predictors of continuous confidence ratings. This is consistent with the results obtained from the ANOVA and, moreover, is consistent with the view that binary confidence scales are more sensitive to low levels of awareness than are continuous scales.

## DISCUSSION

The data reported here replicate an earlier experiment by Vokey and Brooks (1992) and show that grammaticality decisions are based on both the grammatical status of items and their similarity to study items. The question of whether learning of this kind involves awareness has aroused considerable controversy. Subjective reports of confidence are frequently used to assess awareness, and the purpose of the experiment reported here was to determine whether different influences on categorization—namely, grammaticality and similarity—influence confidence ratings.

The key finding is that confidence ratings made on the continuous scale that has most often been used to assess levels of awareness are closely related to the grammaticality of test items but are unrelated to similarity. If taken at face value, the data seem to support the view that rule-based categorization involves explicit knowledge and similarity-based categorization involves implicit knowl-

edge. However, confidence ratings made on a binary scale were influenced by both grammaticality and similarity, indicating that both influences on categorization are available to awareness.

One possibility is that continuous confidence ratings are more closely related to a different index of similarity than the one used here. Indeed, the measure of similarity used here has been criticized by a number of researchers (e.g., Perruchet, 1994) because it assumes that participants memorize whole exemplars and classify new items on the basis of their similarity to specific study items. Perruchet suggested, instead, that the effects of both similarity and grammaticality could be explained by a single fragment-based theory in which participants memorize small fragments of study items and classify test items on the basis of how many memorized fragments they contain. Thus, each test item's similarity to all of the study exemplars can be expressed numerically as chunk strength. Although there has been much debate as to what is the most appropriate measure of similarity, these kind of measures strongly predict the probability that a particular item will be endorsed or rejected (Johnstone & Shanks, 1999). This model is intuitively appealing because of its parsimony (being based on a single process) and because it is similar to neural network simulations of artificial grammar learning (Dienes, 1992). Some proponents of this model have argued that memory for fragments is explicit, because it is available to direct tests of knowledge (Dulany, Carlson, & Dewey, 1984; Perruchet & Pacteau, 1990). By contrast, others have suggested that the processes involved in artificial grammar learning are akin to recollection and familiarity-based memory as they are formulated in dual-process theories of recognition (e.g., Higham, Vokey, & Pritchard, 2000). According to this model, recollection involves explicit memory for specific exemplars, whereas familiarity corresponds to similarity as it is formulated in fragment models. However, in this case, fragments are thought to be stored implicitly. The data reported here support the former position, in the sense that the measures of similarity based on chunk strength reliably predicted binary confidence ratings, indicating that these influences were available to awareness. Moreover, even continuous con-

**Table 2**  
Standardized Regression Coefficients for Each Similarity Index by Type of Confidence Rating

	Continuous Confidence				Binary Confidence			
	$\beta$	SE	<i>t</i>	<i>df</i>	$\beta$	SE	<i>t</i>	<i>df</i>
Bigram chunk strength								
Endorsements	0.05	0.03	1.56	40	0.09**	0.03	2.70	40
Rejections	-0.08*	0.04	-2.05	41	-0.13	0.03	-3.91	35
First and last letter strength								
Endorsements	0.11**	0.03	3.15	40	0.09**	0.03	2.75	40
Rejections	-0.05	0.04	-1.29	41	-0.09**	0.03	-3.03	35
Anchor position chunk strength								
Endorsements	-0.01	0.03	-0.48	40	-0.08	0.03	-0.24	40
Rejections	-0.10**	0.03	-3.49	41	-0.10**	0.03	-3.67	35

\*Significant at .05 level. \*\*Significant at .01 level.

confidence ratings were predicted by some, but not all, measures of chunk strength.

The pattern of data supports previous findings that binary confidence scales are more sensitive to low levels of awareness than continuous scales are (Tunney & Shanks, 2003). There exists no satisfactory explanation of why the two scales differ in sensitivity, despite the fact that the difference has been observed several times. Kunitomo et al. used signal detection theory to analyze binary confidence ratings in their experiments, in effect applying a zero-correlation criterion, and a guessing criterion to assess continuous ratings. They concluded that the guessing criterion applied to the continuous scale was contaminated by response bias, whereas their signal detection analysis of the binary scale separated bias from sensitivity. However, Tunney and Shanks found that even when the continuous and the binary scales were analyzed in the same way, the binary scales remained more sensitive than the continuous scales. They concluded that participants simply find it easier to place their phenomenal states into binary categories than onto continuous scales. The experiment reported here indicates that all of the information that it has been suggested that people acquire in artificial grammar learning is, to a measurable extent, available to awareness.

#### REFERENCES

- CHAN, C. (1992). *Implicit cognitive processes: Theoretical issues and applications in computer systems design*. Unpublished doctoral thesis, University of Oxford.
- CHEESMAN, J., & MERIKLE, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics*, **36**, 387-395.
- DIENES, Z. (1992). Connectionist and memory-array models of artificial grammar learning. *Cognitive Science*, **16**, 41-79.
- DIENES, Z., & ALTMANN, G. T. M. (1997). Transfer of implicit knowledge across domains: How implicit and how abstract? In D. C. Berry (Ed.), *How implicit is implicit learning?* (pp. 107-123). Oxford: Oxford University Press.
- DIENES, Z., ALTMANN, G. T. M., KWAN, L., & GOODE, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1322-1338.
- DIENES, Z., & BERRY, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review*, **4**, 3-23.
- 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.
- ERIKSEN, C. W. (1960). Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review*, **67**, 279-300.
- HIGHAM, P. A. (1997a). 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. (1997b). Dissociations of grammaticality and specific similarity effects in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 1029-1045.
- HIGHAM, P. A., VOKEY, J. R., & PRITCHARD, J. L. (2000). Beyond dissociation logic: Evidence for controlled and automatic processing in artificial grammar learning. *Journal of Experimental Psychology: General*, **129**, 457-470.
- 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, & Cognition*, **25**, 524-531.
- JOHNSTONE, T., & SHANKS, D. R. (2001). Abstractionist and processing accounts of implicit learning. *Cognitive Psychology*, **42**, 61-112.
- KINDER, A., & SHANKS, D. R. S. (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., RAMUS, S. 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, & Cognition*, **20**, 79-91.
- KUNIMOTO, C., MILLER, J., & PASHLER, H. (2001). Confidence and accuracy of near-threshold discrimination responses. *Consciousness & Cognition*, **10**, 294-340.
- 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. (1997). Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 1007-1028.
- PEIRCE, C. S., & JASTROW, J. (1884). On small differences of sensation. *Memoirs of the National Academy of Sciences*, **3**, 75-83.
- PERRUCHET, P. (1994). Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks (1992). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 223-228.
- PERRUCHET, P., & PACTEAU, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, **119**, 264-275.
- POTHOS, E. M., & BAILEY, T. M. (2000). The role of similarity in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 847-862.
- REBER, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning & Verbal Behavior*, **6**, 855-863.
- SHANKS, D. R., & ST. JOHN, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral & Brain Sciences*, **17**, 367-447.
- TUNNEY, R. J., & ALTMANN, G. T. M. (1999). The transfer effect in artificial grammar learning: Reappraising the evidence on the transfer of sequential dependencies. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 1322-1333.
- TUNNEY, R. J., & SHANKS, D. R. (2003). Subjective measures of awareness and implicit cognition. *Memory & Cognition*, **31**, 1060-1071.
- 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.
- YONELINAS, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, **25**, 747-763.

**APPENDIX**  
**Stimuli Used in Experiment 1**

Training Items (Grammatical)	Test Items	
	Grammatical	Ungrammatical
	List 1	
MXRVXT	MXRMXT	MXRRXT
VMTRRRR	VMTRRRX	VMTRRRT
MXTRRR	VXTRRR	TXTRRR
VXVRMXT	VXVRVXT	VXVRTXT
VXVRVM	VXVRVV	VXVRVT
VMRVVVV	VMRVVVM	VMRVVVR
MXRTMVR	MXRTMXR	MXRTMTR
VMRMXTR	VMRVXTR	VMRTXTR
MXR	MVR	MTR
VMRVXVR	VMRVXVT	VMRVXXV
MVRVM	MXRVM	MTRVM
VMRMVRV	VMRMXRV	VMRMTRV
VMRMVXR	VMRMVXT	VMRMVXX
MXRTVXT	MXRTMXT	MXRTRXT
MXRMVXR	MXRMVXT	MXRMVXX
MVXTR	MVXTX	MVXTT
	List 2	
MVXRM	MVXRV	MVXRT
MVXRMVR	MVXRMXR	MVXRMTR
VXVT	VXVR	VXVM
MXRMXRV	MXRMVRV	MXRMTRV
MXRTVMT	MXRTVMR	MXRTVMM
VMT	VMR	VMM
MVXRVMR	MVXRVMT	MVXRVMX
MVXRVVV	MVXRVVM	MVXRVVR
MXRMXRM	MXRMVRM	MXRMTRM
VMRVMT	VMRVMR	VMRVMM
MXRVVVM	MVRVVVM	MTRVVVM
VXVRMXR	VXVRMVR	VXVRMTR
MXTRRRX	VXTRRRX	TXTRRRX
VXVTRRX	VXVTRRR	VXVTRRM
MXTRRX	VXTRRX	TXTRRX
VMRTMXT	VMRTVXT	VMRTTXX

(Manuscript received August 28, 2003;  
revision accepted for publication May 6, 2004.)