Journal of Experimental Psychology:
Learning, Memory, and Cognition
1993, Vol. 19, No. 3, 561–581

0278-7393/93/$3.00

Memory-Based Automaticity in the Discrimination of Visual Numerosity

Mary E. Lassaline and Gordon D. Logan

In the development of memory-based models of automaticity, it is crucial to specify the nature of the memory representation. Seven experiments with 94 students use a counting task to determine whether a feature (i.e., identity, color, or orientation) is explicitly represented in memory. It is assumed that the degree of transfer to a pattern differing on one feature is determined by that feature's importance in supporting skilled performance. Experiment 1 determined the practice necessary to obtain automaticity. In Experiments 2a, 3a, and 4a, which investigated the nature of the representation after extended practice, changing neither the identity nor color of elements had strong effects on transfer, but changing pattern orientation did impair memory retrieval, thus suggesting that for the counting task, pattern orientation is more important than element identity or color. Experiments 2b, 3b, and 4b replicated these results after limited practice.

Memory-based models explain automaticity as a qualitative change in performance from algorithmic computation to memory retrieval. Logan's (1988a, 1988b, 1990, 1992) instance theory is a special case of memory-based models that characterizes processing as a race between the algorithm and memory retrieval. In this article, instance theory is extended to account for the development and transfer of automaticity with nonsymbolic stimuli. More specifically, item specificity of transfer, predicted by memory-based theories, is tested. Also, the nature of the memory representation of an instance is investigated.

Instance theory relates automaticity to the acquisition of a domain-specific knowledge base, rather than the gradual decrease in consumption of some limited capacity. Each time a task is performed, a separate episodic trace (an *instance*) is stored in memory. With practice, more instances accumulate in memory, and thus it becomes more probable that the solution will be retrieved from memory before it is produced by the algorithm. Performance is automatic when it is based on single-step, direct-access memory retrieval rather than algorithmic computation.

Transfer of memory-based automaticity should be narrow, because learning is tied to specific examples studied during training. Beyond making explicit predictions regarding transfer of learning, memory-based automaticity leads to an interest in the underlying representation, which is deemphasized or ignored in other approaches to automaticity. Without understanding which aspects of the external stimulus are preserved in the memory representation, it is difficult to test the predictions made by instance theory in particular and by memory-based models in general (or any model that relies on a notion of transfer in making predictions). How discrepant can a transfer item be from a learning item while still supporting memory retrieval? What exactly qualifies as narrow transfer?

In contrast to memory-based models, process-based explanations of automaticity (Anderson, 1982; Logan, 1978; Shiffrin & Schneider, 1977) suggest that it is the execution of a process that improves with practice and thus is responsible for improvements in performance, rather than the accumulation of instances. Process-based explanations, therefore, do not predict item-specific learning or narrow transfer. As an alternative to a purely memory-based or process-based account, a combination of these two classes of models would offer an account of automaticity under which memory retrieval supports an improvement in a process by eliminating inefficient variations of that process (suggested in Logan, 1988a). This possibility, not clearly articulated in the literature, will be discussed further here.

Although most of the research supporting process-based automaticity uses simple stimuli in a visual search task (Schneider & Fisk, 1984; Schneider & Shiffrin, 1977; Shiffrin & Dumais, 1981; Shiffrin & Schneider, 1977), previous research exploring instance theory has used linguistic stimuli in symbolic tasks such as lexical decision and alphabet arithmetic (Logan 1988a, 1990; Logan & Klapp, 1991). Both of these cases provide strong support for instance theory: The quantitative predictions described earlier were realized, and item-specific learning was demonstrated. However, the extent to which instance theory can account for nonverbal processes operating on nonlinguistic stimuli is as yet unclear. An ideal paradigm would involve a nonlinguistic task, one that is perhaps easier to understand than visual search.

A counting task, in which subjects are presented with visual patterns of discrete elements and asked to determine the numerosity of the elements, fits the above prescription, because it provides a way to separate memory-based and process-based automaticity on their predictions regarding item-specific learning and transfer. Perhaps more important, the counting task can be described by a simple algorithm that

Mary E. Lassaline (now at Department of Psychology, University of Michigan) and Gordon D. Logan, Department of Psychology, University of Illinois.

This research was supported by National Science Foundation Grant BNS 88-10365 to Gordon D. Logan and was conducted while Mary E. Lassaline held a Dallenbach Fellowship at the University of Illinois.

Correspondence concerning this article should be addressed to Gordon D. Logan, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, Illinois 61820.

can, in principle, be replaced by memory retrieval. From the perspective of instance theory, it is important that counting provides a clear, consensually agreed upon algorithm that predicts the linear increase in reaction time (RT) with an increase in elements to be counted. The counting algorithm is distinct from the memory retrieval process proposed by instance theory. Thus, distinguishing algorithmic computation from memory retrieval should be relatively easy.

The counting algorithm involves three component processes: (a) indexing the stimulus elements, (b) mapping a number from the number line to each element, beginning with 1 and ending with the number mapped to the last element in the set, and (c) producing a numerosity response consisting of the last number mapped. Counting initially requires serial spatial indexing and mapping (i.e., Pylyshyn, 1989), in which a set of visible objects is placed in one-to-one correspondence with an internal representation of the number line. Each element of a display requires some amount of time to be processed; therefore, time to produce a numerosity response by spatial indexing should increase monotonically with the number of elements in a display. Previous research has shown that the increase is linear, with a mean slope of about 300 ms per item for displays of from 4 to 10 items (Chi & Klahr, 1975).

Memory-based theories suggest that with repetition of a particular pattern and practice at determining its numerosity, presentation of that pattern should act as a cue for retrieval of the correct number from memory. With practice, the stimulus evokes memory retrieval, without counting. As more traces of a pattern and its corresponding numerosity accrue in memory, it becomes more probable that, upon presentation of the pattern, retrieval of its numerosity will occur before the counting algorithm has produced a solution. In this case, the slope of the function relating response latency to the number of elements in the display should flatten to an asymptotic level. If memory retrieval does preempt algorithmic computation in producing a response, and if visual stimulus processing and retrieval time are both comparable across level of numerosity (taking no more time to visually process or to retrieve the numerosity of an 11-element display than to visually process or to retrieve the numerosity of a 6-element display), then the slope of the function relating RT to the number of elements should decrease to zero over training blocks. It is not likely that retrieval time will vary with numerosity, but it is conceivable that as complexity of the visual display increases with numerosity, it may take longer to simply encode, at a very low level, the visual display. This would explain a reduction in slope to a nonzero number.

Memory-based models predict that with presentation of novel exemplars, any such practice effect should be eliminated and performance should regress back to an earlier level, because learning is item specific. The nature of the memory representation becomes important in the determination of just what qualifies as a novel exemplar. In our experiments, we used the following logic to investigate the nature of the representation. First, an attribute of the stimuli used during training (e.g., the shape of the elements composing each stimulus pattern) is isolated. The value of that attribute is changed (e.g., from circular to square) to create a novel set of stimuli that preserves everything but the isolated attribute (shape) of the training stimuli. If skilled performance in the task is disrupted such that it regresses to an unskilled level when the novel stimuli are presented in the transfer test, it can be assumed that the isolated attribute was somehow involved in the representation of the original training stimuli in memory. If performance with the novel stimuli remains at a practiced level, then there is no evidence that the isolated attribute is represented in the memory that supports skilled performance in this particular task. The importance of the attribute in the memory that supports skilled performance determines the degree of transfer to the novel stimuli. It is important to point out, though, that such a pattern of results is bound to the specific task or skill (numerosity judgments) and to the specific test of memory (transfer of skilled performance) used. The use of a different task (shape judgments) or memory test (old/new recognition) may paint a different picture of the representation of a training stimulus. Given that we are primarily interested in the memory that supports skilled performance, transfer of skill to performance with novel stimuli is the appropriate measure.

Our experiments used displays of 6 to 11 elements to eliminate the possibility of performing a counting task by subitizing (directly apprehending numerosity without counting). The upper limit of subitizing is typically found to be approximately 5 elements (Jensen, Reese, & Reese, 1950; Kaufman, Lord, Reese, & Volkmann, 1949; Mandler & Shebo, 1982). It is possible that subpatterns within a pattern may be subitized, for example, the numerosity of a 12-element display could be determined by directly perceiving triads, and counting four such triads. Such a process may require less time than serially counting all 12 elements. Nevertheless, the number of groups to be counted should increase with numerosity, so RT should still increase (relatively linearly) with numerosity. We assume that the only ways to produce a response in the counting task are counting, remembering, and guessing. Guessing is reduced or eliminated by requiring a high degree of accuracy.

The experiments described below used this counting task. Experiment 1 demonstrated that automaticity can be developed by extended practice at the counting task. Experiments 2–4 tested the item specificity of learning by presenting old and new items in a transfer test. In addition, these experiments addressed the representation issue by determining whether some feature of the visual pattern is explicitly represented in memory. Experiments 2a, 3a, and 4a examined the nature of the memory representation across extended practice. Experiments 2b, 3b, and 4b examined the nature of the memory representation at an intermediate stage of practice. In principle, instance theory predicts that the same results should be obtained regardless of the level of practice. We included two levels of practice in these experiments so that we could test that hypothesis and generalize our conclusions to a range of practice levels. Three features of the visual pattern that were investigated include the identity of each element of the display (Experiments 2a and 2b), the organization of the pattern into subgroups (Experiments 3a and 3b), and the spatial configuration of the pattern in the display screen (Experiments 4a and 4b).

Experiment 1

If automaticity is obtained with practice in the counting task, performance should become based entirely on memory retrieval. It should take no longer to respond to a pattern containing 11 elements than to a pattern containing 6, aside from any differences in the initial visual encoding and the response execution component of the RT, and thus the slopes of the function relating response latency to numerosity should be flat. Even if these slopes never completely flatten, they should reach asymptote at some point. Experiment 1 was conducted to determine the asymptotic level of performance.[1] Subjects completed 12 sessions of the counting task, using the same set of stimulus patterns across the 12 sessions, and then were transferred to novel patterns on a final session. Twelve sessions represent 192 exposures to each stimulus, which should be sufficient to produce automaticity (Logan, 1988a). Transfer to new items on Session 13 tests the specificity of what was learned during automatization.

Method

Subjects. Four University of Illinois graduate students from the Psychology Department participated in 13 sessions of this experiment. They each received $50 for their participation.

Apparatus and stimuli. For each subject, 5 unique patterns were generated for each level of numerosity (6, 7, 8, 9, 10, and 11), which led to a total of 30 unique random patterns per subject. Patterns were considered unique when they differed in the position of at least one element. For each pattern, the appropriate number of asterisks was randomly positioned in a 7 × 7 matrix, in the center of an Amdek 722 color monitor. The matrix was not shown on the screen; only the asterisks were displayed. Four sets of patterns were constructed. One set was used for each subject's 12 training sessions. Each set of 30 training patterns was assigned to a different subject to be used in the single transfer session. The constituent elements of each pattern, each 0.3 cm × 0.3 cm, were separated in width by at least 1 cm and in height by at least 1.4 cm, which produced an 8.1 cm × 10.5 cm matrix. IBM-AT personal computers were used to display the stimulus patterns, provide a numberpad on which to make numerosity responses, and collect responses and latencies. The keys on the numberpad marked 1, 2, 3, 4, 5, and 6 were designated to make the numerosity responses "six", "seven", "eight", "nine", "ten" or "eleven". (The numberpad was organized in three rows of three columns, with the 7, 8, and 9 keys on the top row, the 4, 5, and 6 keys in the middle row, and the 1, 2, and 3 keys on the bottom row.)

Procedure. To familiarize subjects with the mapping between numerosity response (6–11) and appropriate response key (1–6), we conducted a series of 60 practice trials. In each practice trial, a fixation point was presented in the center of the screen for 500 ms, followed by a single number selected at random from the range to be used in the experiment. The number remained on the screen until the subject made a response. The interval from the response to the onset of the successive trial was 1,500 ms. Each of the six numbers was presented 10 times during the course of the practice segment, which took approximately 5 min.

Following the practice trials, subjects completed the experimental trials. In each experimental trial, a fixation point was presented in the center of the screen for 500 ms and was followed by a random pattern of asterisks. The pattern remained on the display screen until a response was made on the numberpad; again, subjects were instructed to respond as accurately and as rapidly as possible. This was done to prevent guessing and to eliminate estimation of numerosity as an alternative means of performing the task (see Mandler & Shebo, 1982). After making sure subjects clearly understood the task, we gave them four blocks of 120 training trials, and each block was followed by an optional break. Trials were blocked such that no pattern was repeated until all other patterns had been presented and were randomized within blocks. Before the practice trials, and again before the experimental trials, subjects were instructed to respond with the appropriate number as rapidly and as accurately as possible. Also, subjects were instructed to try not to look at the response keys, because the number appearing on the key (1–6) was inconsistent with the response (6–11).

In each session, each of the five training exemplars generated for each level of numerosity (6–11) was presented four times in each of four 120-trial blocks, for a total of 16 presentations per item. There were a total of 480 trials per session, which took approximately 40 min to complete.

Each subject completed 13 sessions of the counting task, and each session consisted of 60 trials of practice at making the keypad response to a displayed number followed by the 480 experimental trials. The first 12 sessions used a single set of 30 patterns for each subject, such that each subject received a total of 5,760 training trials, 192 trials of each training pattern. During Session 13 a different set of patterns was presented in a transfer test, which was preceded by a 60-trial practice phase. The patterns presented to each subject in the transfer test had been presented to a different subject as training patterns, such that each set of patterns was used for both training and transfer.

Results and Discussion

Mean accuracy scores across the 4 subjects for the 12 training sessions were as follows: .95, .96, .96, .97, .97, .97, .93, .96, .97, .96, .98, and .97. Not one of the means across subjects or across sessions was significantly different from another. The mean accuracy for Session 13, during which a set of novel patterns was presented, was .96; again, this was not significantly different from the accuracy of any of the training sessions.

Slopes of the linear function relating response latency to numerosity were calculated for each session after averaging across trial block (1–16). Use of these slopes was justified by a significant linear trend in the response latency data. Two analyses of variance (ANOVAs) were conducted on response latencies: The first examined training response latencies for an effect of numerosity (6–11) and practice (Session 1–12); the second examined transfer response latencies for an effect of numerosity (6–11) and stimulus type (Sessions 1–12 vs. Session 13). The following pattern of results emerged: (a) Response latency increased with numerosity; (b) response latency decreased with training; (c) the increase in response

---

[1] Logan (1992) reported an analysis of the RT distributions from the first 12 sessions of this experiment. He was concerned with changes in the shape of the RT distribution with practice and how those changes related to the shape of the (power function) learning curve. He was not concerned with the transition from counting to remembering, which we address, nor was he concerned with the item specificity of the learning, which is the major focus of this article.

latency with numerosity was smaller as training progressed; (d) response latency increased with numerosity at transfer as well as training; (e) response latency was higher for new transfer patterns than for old patterns; and (f) the increase in response latency with numerosity was smaller for old patterns than for new patterns. This pattern of results is consistent with analogous ANOVAs conducted on the slopes of the function relating response latency to numerosity, which show a decrease in slope with training and an increase in slope to an earlier level of training for new patterns at transfer. Given that both sets of analyses converge, for brevity, only the analyses involving slope are reported here. A table presenting RT as a function of numerosity and practice as well as ANOVA summary tables for the analyses of RT are presented in the Appendix. Goodness of linear fit in the response latency data is assessed in three ways: the proportion of treatment variance accounted for by the linear trend; the squared correlation between mean RT and numerosity; and the square of the correlation between raw RT and numerosity for each subject, averaged across subjects. These statistics are reported in Table 1.

The linear fit is poorer in the latter case than in the second, because in the second case averaging before correlating reduces the variability in response latency across numerosity. In general, the fit is reasonably good, and thus we focus the remainder of the analysis and discussion on slopes. Slopes for the 13 sessions are presented in Figure 1.

Performance reached asymptote after Session 3 following significant improvements in performance over the first two sessions. At transfer, performance regressed to the level displayed during Session 2, which is indicative of item-specific learning. An ANOVA revealed a significant main effect of session, $F(12, 36) = 34.669$, $p < .01$, $MS_e = 1{,}176.039$. A post hoc analysis on the 13 average slopes was conducted, which involved Tukey's honestly significant difference (HSD) test of multiple comparisons and used the studentized range distribution, q (see Neter, Wasserman & Kutner, 1985, pp. 574–584). In this and all following ANOVAs, the main effect of block on slope during training (and of trial type during transfer in Experiments 2–4, which used three types of transfer trials) was analyzed first and then tested post hoc for a linear decrease over trials (and post hoc contrasts were conducted on trial type in transfer for Experiments 2–4). In general, contrasts that did not yield significance post hoc would also have failed to yield significance even under the more powerful a priori method of contrast. Pairwise comparisons revealed no significant differences between Slope 13 (slope from the transfer session) and Slope 2; between Slope 2 and Slope 3; and between all slopes from Sessions 3 to 12. Slopes 1 (initial slope) and 13 (transfer) were significantly different, $q(13, 36) > 6.01$, $p < .01$, $MS_e = 1{,}176.039$, as were Slopes 1 and 2, $q(13, 36) > 5.08$, $p < .05$, $MS_e = 1{,}176.039$. In summary, (a) the initial slope was significantly greater than slopes from all other sessions; (b) the slope from the transfer session was significantly greater than the slopes from Sessions 3 to 12 but was not different from Slope 2; and (c) Slope 2 was significantly greater than the slopes from Sessions 4 to 12 but was not different from Slope 3.

Table 1
Squared Correlations Between Reaction Times and Numerosity for Experiment 1

Session	A	B
1	.88	.82
2	.44	.38
3	.29	.22
4	.10	.07
5	.08	.07
6	.16	.09
7	.18	.11
8	.06	.05
9	.16	.11
10	.16	.08
11	.15	.10
12	.27	.17
13 (Transfer)	.92	.90

Note. Column A = squared correlation between mean reaction times (RT) and numerosity. Column B = average squared correlation between raw RT and numerosity. Proportion of treatment variance accounted for by linear trend in reaction time = .42.

Given that the slope of the linear function relating response latency to numerosity reached asymptote by Session 4, it seems safe to conclude that automaticity in the counting task had been obtained by that point. In accordance with instance theory, performance is automatic when it is based on memory retrieval, and the asymptotic slopes suggest that counting performance was clearly based on memory retrieval after the fourth session. Instance theory also suggests that performance may improve after it becomes automatic (i.e., automaticity is never complete; see Logan, 1985, 1988a) in that memory becomes even more efficient with further practice. Indeed, Logan (1992) analyzed the RT distributions at each numerosity level in each of the 12 training sessions and found substantial improvement in RT after the first four sessions. Logan interpreted this as indicating that memory became stronger with practice even after performance was based entirely on memory retrieval.

In addition, Logan (1992) performed several tests of the distributional assumptions of the instance theory on these data. First, he showed that the whole distribution of RTs decreases as a power function of practice, as instance theory predicts. Second, he fit Weibull distributions to the data and found that the RT distribution had the same shape over sessions, though its scale reduced as a power function of practice, as predicted by instance theory. Third, Logan tested the instance theory prediction that the shape of the retrieval time distribution determines the shape of the learning curve by comparing the exponent of the power function that described the reduction in RT (which determines the shape of the learning curve) to the reciprocal of the exponent of the Weibull distribution (which determines the shape of the retrieval time distribution). He found reasonable agreement between the exponent of the power function and the reciprocal of the exponent of the Weibull distribution. These analyses suggest that instance theory can provide an adequate account of performance changes that occurred during the training phase of this experiment.

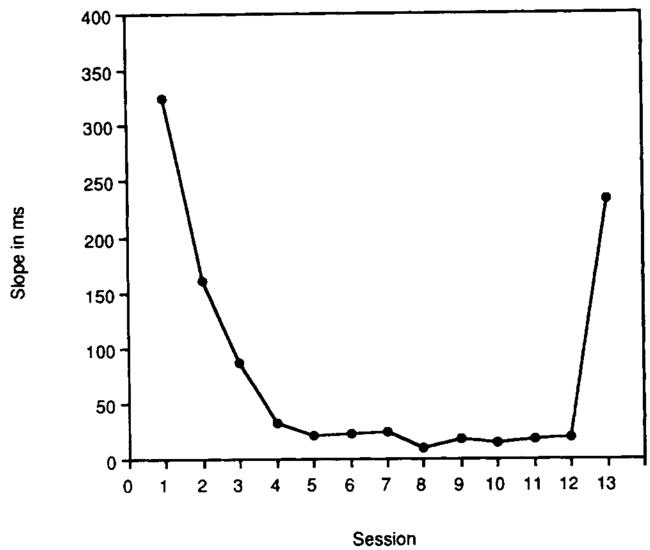

Figure 1. Slope of linear regression function relating reaction times to numerosity as a function of session from Experiment 1. (Sessions 1–12 used a single set of stimulus patterns for each subject; novel patterns were used to test the item specificity of learning in Session 13.)

Regression of transfer performance to the level obtained during early training is particularly strong evidence of item-specific learning given the high degree of practice in this task. The pattern of data from this experiment is consistent with the replacement of the initial algorithm-based performance by performance that is based on memory retrieval over the course of the first several sessions of the experiment. There was a significant change in slope even in the first session, which suggests that automatization was taking place. This provides the motivation for the remaining experiments, directed at clarifying the nature of the process that drives performance after both extended training and the initial session of the counting task, and the nature of the representation on which they operate.

Experiment 2a

Experiment 2a was conducted to test the notion that element identity is preserved in the representation of training patterns after extended training in the counting task. It is possible that as performance becomes driven by instance retrieval rather than by algorithmic processing, the identity of stimulus elements, although unimportant in the counting algorithm, may play a more important role in instance retrieval. This experiment used the same practice and training procedure used in Experiment 1, with the exception that training extended over four sessions rather than 12, as performance in Experiment 1 reached asymptote by the fourth session. Transfer was tested in a final session, as in Experiment 1, but three types of stimulus patterns were presented: patterns used during training (old/old); novel patterns seen only at transfer and constructed using different spatial arrangements of the same set of letters used during training (new/new); and patterns adhering to the same spatial arrangements as those experienced during training but each composed using a different letter from the set used during training (old/new). Refer to Figure 2 for an example of the three stimulus patterns used in this experiment. If element identity is not important for instance retrieval, then performance on old/new transfer patterns should be as good as performance on old/old training items. Alternatively, if element identity plays a role in instance retrieval, then performance on old/new items should be worse than performance on the original training items and possibly as poor as performance with new/new patterns. The degree of importance of element identity for instance retrieval will be reflected in the extent to which performance on old/new items resembles performance on new/new items.

Method

Subjects. Six University of Illinois undergraduates served as subjects. Each subject participated in five 45-min sessions and received $20 for his or her participation.

Apparatus and stimuli. The 90 patterns (30 training patterns, 30 novel transfer patterns, and 30 patterns using the same spatial arrangements as the training patterns but constructed with a different letter) were constructed as in Experiment 1, with the exception that all of the elements within a particular pattern were a single letter drawn from the set A, E, I, O, and U. In each set of 30 patterns (training, novel transfer, and old/new transfer), there was one pattern using each of the five letters at each of the six numerosity levels 6–11. Again, the training and old/new transfer patterns used the same 30 spatial arrangements but a different letter. The *old* in old/new refers to the spatial arrangement of the elements for each old/new pattern; the *new* refers to the single letter assigned to each

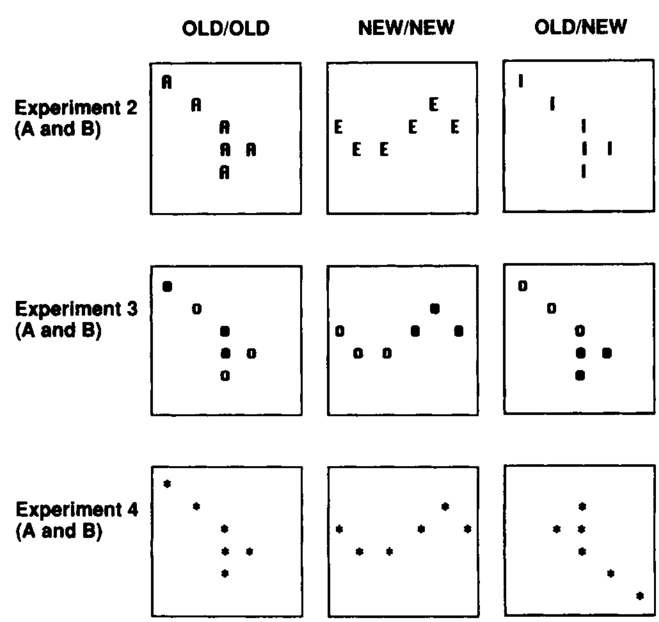

Figure 2. Sample stimulus patterns used in Experiments 2, 3, and 4. (In Experiments 3a and 3b, patterns were composed of red and green asterisks; filled and unfilled letter *o*s are used here for illustrative purposes.)

location within a particular spatial arrangement. The spatial arrangement of the items in each novel transfer pattern was different from that in the training and old/new transfer patterns, and thus the letter assigned to each novel pattern was, by definition, also new to that pattern. An example of the stimulus patterns is displayed in Figure 2. Patterns were displayed, and responses were collected, in the same manner as in Experiment 1.

Procedure. Practice, training, and transfer trials were conducted in the same manner as in Experiment 1, with the exception that training occurred over only 4 sessions rather than 12 (60 practice trials with numbers but no patterns to count, 480 trials of the counting task at each training session), and the transfer session included three trial types (120 trials per each of the following types: old/old, old/new, and new/new).

In each training session, after 60 practice trials (10 at each numerosity level), each of the five training exemplars generated for each level of numerosity was presented four times in each of four 120-trial blocks, for a total of 16 presentations per item. The 60 practice trials and the 480 experimental trials took approximately 45 min to complete.

The final session included four 90-trial blocks, with one presentation of each of the 30 training, novel transfer, and old/new transfer patterns in each block.

Design. Session (1–4), level of numerosity (6–11), and type of transfer pattern (old/old, new/new, and old/new) were manipulated within subject.

Results and Discussion

Mean accuracy across the four training sessions and across the six levels of numerosity was .95. None of the means across subjects, sessions, and numerosity levels was significantly different from another.

Two ANOVAs were conducted on response latencies, as in Experiment 1: The first examined training response latencies for an effect of numerosity (6–11) and practice (Session 1–4); the second examined transfer response latencies for an effect of numerosity (6–11) and stimlus type (old/old, old/new, and new/new). Across training sessions, response latencies decreased with session and increased with numerosity, and the increase in response latency with a greater number of elements per pattern was lessened as training progressed across the four sessions. In other words, the slope of the linear regression function relating response latency to numerosity decreased with training. This pattern of results is consistent with analogous ANOVAs conducted on the slopes of the function relating response latency to numerosity, which show a decrease in slope with training and an increase in slope to an earlier level of training for new patterns at transfer. Given that both sets of analyses converge, for brevity, only the analyses involving slope are reported here. A table presenting RT as a function of numerosity and practice as well as ANOVA summary tables for the analyses of RT are presented in the Appendix. Goodness of linear fit in the response latency data is assessed in three ways: the proportion of treatment variance accounted for by the linear trend; the squared correlation between mean RT and numerosity; and the square of the correlation beween raw RT and numerosity for each subject, averaged across subjects. These statistics are reported in Table 2.

In general, the fit is good; therefore the remaining analyses focus on slope. The average slope as a function of session is presented in Figure 3.

The mean training slope from Session 1 was 373 ms per item, which dropped to 274 ms per item in Session 2, 195 ms per item in Session 3, and 109 ms per item in the final training session. There was a significant effect of session on training slope, $F(3, 15) = 37.87$, $p < .01$, $MS_e = 30{,}106$. A post hoc analysis on training slopes was conducted with the Tukey HSD method of multiple comparisons. Pairwise comparisons revealed significant differences between Slopes 1 and 4, $q(4, 15) > 4.08$, $p < .05$, $MS_e = 30{,}106$, suggesting that performance did improve with training.

At transfer, mean accuracy across numerosity levels (6–11) for the three stimulus types was as follows: old/old, .95; new/new, .95; and old/new, .96. None of the means across stimulus type and numerosity level was statistically different from another. Response latency for old/old patterns was faster than for old/new patterns, which were in turn faster than new/new patterns. Response latency increased with numerosity, and the increase in response latency with numerosity for old/old patterns was smaller than for old/new patterns, which were in turn smaller than new/new patterns, which suggests that slopes as well as response latencies decreased from new/new to old/new and old/old patterns. The slope for old/old patterns was 159 ms per item; for old/new patterns, 213 ms per item; and for new/new patterns, 374 ms per item. There was a significant effect of stimulus type of slope, $F(2, 10) = 12.04$, $p < .05$, $MS_e = 6{,}294$. Post hoc contrasts using the Tukey HSD method revealed that slopes for old/old and old/new patterns were not significantly different from each other, $q(3, 10) < 3.88$, $p > .10$, $MS_e = 6{,}294$, whereas the slopes of both types of old stimulus patterns (old/old and old/new) were significantly different from that for new/new patterns, $q(3, 10) > 3.88$, $p < .05$, $MS_e = 6{,}294$.

The results of this study are consistent with results from Experiment 1. Learning was item specific, because the improvement in performance obtained across training did not

Table 2
Squared Correlations Between Reaction Times and Numerosity for Experiments 2a, 3a, and 4a

	Experiment 2a		Experiment 3a		Experiment 4a	
Session	A	B	A	B	A	B
1	.94	.40	.95	.24	.96	.31
2	.88	.27	.86	.10	.87	.18
3	.84	.17	.86	.05	.83	.15
4	.54	.09	.68	.02	.67	.07
5 (Old/old)	.77	.12	.58	.04	.80	.17
5 (New/new)	.91	.36	.85	.18	.98	.28
5 (Old/new)	.94	.10	.80	.06	.96	.20

Note. Column A = squared correlation between mean reaction times (RT) and numerosity. Column B = average squared correlation between raw RT and numerosity. The proportion of treatment variance accounted for by linear trend was .28, .79, and .69 for Experiments 2a, 3a, and 4a, respectively.

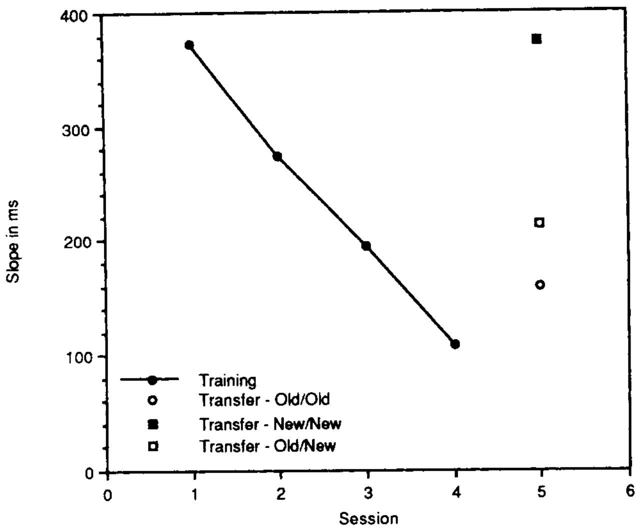

Figure 3. Slope of function relating reaction time to numerosity as a function of training session from Experiment 2a. (Sessions 1–4 used a single set of stimulus patterns for each subject; novel patterns were used to test the item specificity of learning in Session 5.)

transfer to novel patterns. In addition, there was significant transfer to old/new patterns, which suggests that element identity is not important for item-specific learning in this task. The next experiment tests the importance of element identity for item-specific learning at an intermediate stage of practice. Again, the degree of importance of element identity for instance retrieval is reflected in the extent to which performance on old/new items resembles performance on new/new items.

Experiment 2b

The results of Experiment 2a suggest that element identity is not a very important part of the representation of the training patterns after extended training in the counting task. In that experiment, performance on old/new transfer items (patterns adhering to the same spatial arrangements as those experienced during training but composed of novel letters) was almost as good as performance on old/old (training) items. Experiment 2b was conducted to test memory representation for the inclusion of element identity information at an intermediate stage of practice. If element identity is not important for item-specific learning, then performance on old/new transfer items should be as good as on training items (old/old). If, in contrast, element identity is important for item-specific learning, then performance on old/new transfer items should be worse than that on the original training items and possibly as poor as performance with novel patterns constructed from different letters and different spatial arrangements than those used during training (new/new patterns). Old/new performance falling between old/old and new/new performance indicates that element identity has some importance for instance retrieval.

Method

Subjects. Twenty-four University of Illinois undergraduates served as subjects. Each subject participated in a 1-hr session, either for pay or for credit in an introductory psychology course. The data from 2 additional subjects were not used in any analyses because of experimenter error.

Apparatus and stimuli. The 90 patterns (30 training, 30 novel transfer, and 30 old/new transfer patterns) were constructed in the same manner as in Experiment 2a, with the exception that letters were drawn from the set [B, C, D, F, G, H, J, K, L, M, N, P, Q, R, S]. Five letters were randomly selected from this set for use in the training patterns; five different letters were selected for use in the novel transfer patterns; and for the old/new patterns, five letters that had not been used in the training patterns or in the novel transfer patterns were used, but these were arranged spatially in the same configuration as in the 30 original training patterns. The *old* in old/new refers to the spatial arrangement of the elements in the pattern; the *new* refers to the letter that was used to construct the pattern. The new/new patterns differed from the new/new patterns used in Experiment 2a, because the letters used in the new/new patterns in this experiment had not been used in either the old/old or old/new patterns. This difference is not relevant to the numerosity judgments made in each experiment, because in both experiments, each letter, old and new, was used at each numerosity level. In all, subjects saw 90 different patterns (30 throughout the 16 training blocks and at transfer, the same 30 with a different letter, and 30 new patterns only at transfer). Patterns were displayed, and responses were collected in the same manner as described in Experiments 1 and 2a.

Procedure. Practice, training, and transfer trials were conducted in the same manner as in the previous experiments (60 practice trials with numbers but no patterns to count and 480 trials of the counting task), with the exception that 90 trials were included in the transfer test at the end of the single session. There was one presentation of each of three types of transfer trials, with 30 patterns of each type. With training and transfer, subjects completed 570 trials (rather than 480 training trials at each session).

Design. Training block (1–16), level of numerosity (6–11), and type of transfer pattern (training pattern, old/old; novel pattern seen only at transfer, new/new; and the original training patterns with a new color configuration, old/new) were manipulated within subject.

Results and Discussion

The mean accuracy across the 16 training blocks was .93. At transfer, the mean accuracy for each type of pattern was as follows: old/old, .94; new/new, .94; and old/new, .94. There was no effect of practice or type of transfer trial on accuracy.

Two ANOVAs were conducted on response latencies, as in the multisession experiments: The first examined training response latencies for an effect of numerosity (6–11) and practice (Training Blocks 1–16); the second examined transfer response latencies for an effect of numerosity (6–11) and stimulus type (old/old, old/new, and new/new). The same pattern of response latency results described for previous multisession experiments was obtained in this single-session experiment, which justified an analysis of the slopes of linear regression functions relating RT to numerosity. This pattern of results is consistent with analogous ANOVAs conducted on the slopes of the function relating response latency to

Table 3
Squared Correlations Between Reaction Times and Numerosity for Experiments 2b, 3b, and 4b

Session	Experiment 2b		Experiment 3b		Experiment 4b	
	A	B	A	B	A	B
1	.84	.29	.87	.36	.93	.32
2	.93	.42	.87	.38	.89	.36
3	.96	.44	.82	.35	.96	.40
4	.92	.40	.94	.42	.90	.40
5	.91	.39	.94	.39	.91	.37
6	.91	.39	.89	.41	.94	.36
7	.92	.36	.92	.37	.96	.36
8	.91	.37	.93	.38	.85	.33
9	.95	.34	.95	.34	.91	.31
10	.91	.35	.93	.40	.88	.29
11	.82	.28	.94	.36	.88	.26
12	.86	.31	.93	.38	.85	.28
13	.86	.28	.90	.30	.87	.26
14	.91	.29	.94	.30	.90	.26
15	.79	.18	.92	.38	.90	.26
16	.85	.26	.88	.30	.90	.28
Old/old	.86	.21	.96	.29	.74	.20
New/new	.95	.36	.94	.31	.92	.32
Old/new	.86	.29	.90	.33	.89	.33

Note. Column A = squared correlation between mean reaction time (RT) and numerosity. Column B = average squared correlation between raw RT and numerosity. The proportion of treatment variance accounted for by linear trend was .91, .93, and .92 for Experiments 2b, 3b, and 4b, respectively.

numerosity, which show a decrease in slope with training and an increase in slope to an earlier level of training for new patterns at transfer. Given that both sets of analyses converge, for brevity, only the analyses involving slope are reported here. A table presenting RT as a function of numerosity and practice as well as ANOVA summary tables for the analyses of RT are presented in the Appendix. Goodness of linear fit in the response latency data is assessed in three ways like it was in Experiments 1, 2a, and 2b: the proportion of treatment variance accounted for by the linear trend; the squared correlation between mean RT and numerosity; and the square of the correlation between raw RT and numerosity for each subject, averaged across subjects. These statistics are reported in Table 3.

The linear fit is poorer in the latter case than in the second, because in the second case, averaging before correlating reduces the variability in response latency across numerosity. In general, the fit is reasonably good, and so we focus the remainder of the analysis and discussion on slopes.

Unlike previous multisession experiments, two sets of slopes were calculated and analyzed: The first included mean RTs from all six numerosity levels (6–11), and the second dropped RTs from the endpoints (6 and 11). Analyses using the second set of slopes (of lines fit to mean RTs at 7, 8, 9, and 10 elements) were motivated by two concerns: First, patterns at the endpoints of the numerosity levels used in the present experiments may be more susceptible to strategy use during the single session, during which complete automaticity did not obtain; and second, data from the initial practice at making the keypad response to each number showed that response times to the numbers 6 and 11 were faster than response times to the remaining four numbers.

The results from both sets of analyses converged. The average slopes across training block and at transfer from the first set (including mean RTs from all six numerosity levels) are presented in Figure 4. In the initial training block, the mean slope was 316 ms per item. The mean slope increased over subsequent training blocks, then decreased, dropping to 276 ms per item in the final block. There was a significant main effect of training block, $F(15, 345) = 3.946, p < .001, MS_e = 10{,}499$, and a significant linear decrease over training blocks, $F(1, 23) = 14.178, p < .01, MS_e = 27{,}880$, which suggests that performance did improve with practice. The analysis of the second set of slopes, those calculated after dropping the RTs from the first and last numerosity levels, showed no effect of training block, $F(15, 345) = 1.340, p = .18, MS_e = 25{,}396$, but the linear decrease over training blocks did approach significance, $F(1, 23) = 3.663, p = .07, MS_e = 71{,}036$.

At transfer, the slope of the function relating RT to numerosity (including RTs from all numerosity levels) was 266 ms per item for old/old patterns, 344 ms per item for new/new patterns, and 285 ms per item for old/new patterns. The ANOVA conducted on these transfer slopes revealed a main effect of type of transfer pattern, $F(2, 46) = 4.946, p < .05, MS_e = 8{,}054$. Post hoc Tukey contrasts showed a significant difference between both types of old patterns (old/old and old/new) and the new patterns (new/new), $q(3, 46) > 3.44, p < .05, MS_e = 8{,}054$, but no difference between the two types of old patterns, $q(3, 46) < 3.44, p > .05, MS_e = 8{,}054$. Two further contrasts—between initial slope at the beginning of training and slope for new patterns at transfer (316 ms per item and 344 ms per item, respectively) and between final training slope and slope for old/old patterns at transfer (276

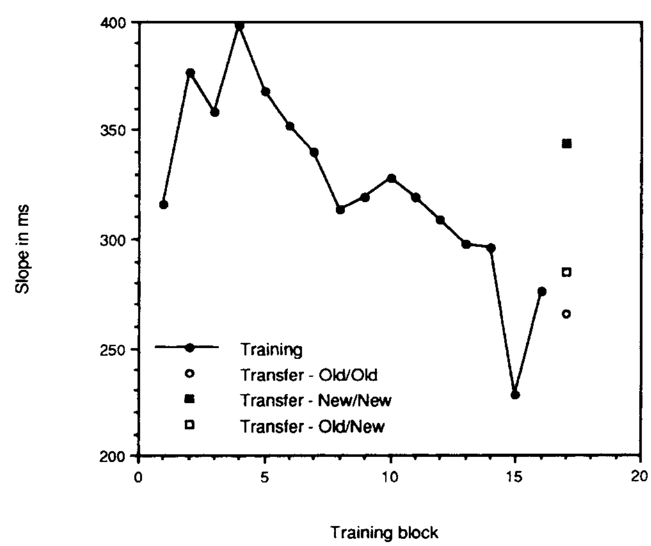

Figure 4. Slope of function relating reaction time to numerosity as a function of training block (1–17) from Experiment 2b. (New/new and old/new items were included in Block 17 for a transfer test.)

ms per item and 266 ms per item, respectively)—were not significant.

The transfer slopes calculated after dropping RTs from the endpoints of the range of numerosity levels used in this task were 257 ms per item for old/old patterns, 356 ms per item for new/new patterns, and 276 ms per item for old/new patterns. These average slopes are very similar to those obtained when including RTs from all numerosity levels (266, 344, and 285 ms per item, respectively), yet analysis of transfer data from this second set of slopes revealed no main effect of transfer type, $F(2, 46) = 1.353$, $p = .27$, $MS_e = 48{,}714$. Because the mean slopes from this second set are comparable to those reported from the first set calculated using all numerosity levels, it is likely that the differences between old and new transfer patterns would approach significance with an increase in power. As in the analysis of the first set of slopes, there was no difference between initial slope and slope for new/new patterns at transfer (323 ms per item and 356 ms per item, respectively) nor was there a difference between final training slope and slope for old/old patterns at transfer (280 ms per item and 257 ms per item, respectively).

These results suggest that learning was item specific, because the improvement in performance obtained across training did not transfer to novel patterns. Also, good transfer to old/new patterns, which were composed of letters different from those experienced during training, confirmed the notion suggested in Experiment 2a that element identity is not important for item-specific learning in this task. The next two experiments test the notion that the organization of subgroups within a pattern is preserved in memory, along with the location of individual elements.

Experiment 3a

Is the organization of smaller patterns within each pattern preserved in an instance? Experiment 3a examined the role of subgroup configuration in representation by maintaining the overall pattern and identity of training exemplars at transfer but changing their constituent structure. This was accomplished using color to create perceptual groups. Grouping by color is an example of grouping by similarity, an idea dating back at least to the Gestalt psychologists (for a review of the Gestalt movement see Boring, 1950). Previous research (Kahneman & Henik, 1981; Treisman & Gelade, 1980) suggested that color determines the perceptual organization of visual patterns. This experiment capitalized on this finding by training subjects on patterns composed of red and green elements, with grouping by color providing an intermediate level of structure to each pattern. This intermediate structure was then changed at transfer by changing the color configuration of the training patterns.

Each pattern was composed of a number of asterisks randomly colored either red or green, such that half of the components were one color and half were the other. As in our previous experiments, subjects were repeatedly presented with a number of such patterns, and we expected the improvements in performance described earlier.

On transfer to new patterns, memory-based theories predicted that performance would return to an earlier unpracticed level. In addition, a third type of transfer pattern was included in this experiment. The color of the elements of each training pattern was randomly reassigned (red or green) with the constraint that the color reassignment not be identical to the original color assignment, which thereby changed the color configuration of the pattern while preserving its spatial configuration. This experiment used the same practice, training, and transfer procedure that was used in Experiment 2a. Again, as in Experiment 2a, transfer was tested in a final session but with the following three types of stimulus patterns: patterns used during training (old/old); novel patterns seen only at transfer, which were constructed by using different spatial arrangements of asterisks (new/new); and patterns that adhered to the same spatial arrangements as those experienced during training but were composed of a different color configuration than the set used during training (old/new). Refer to Figure 2 for an example of the three stimulus patterns used in this experiment. If the color configuration is preserved in memory, then performance on these old patterns with new color configuration should regress to an unpracticed level, along with that of the new transfer patterns. If the representation is organized in terms of groups, then a change in the constituent structure of the grouping should impair memory retrieval, which may be evidenced by poor performance on the old patterns with a different color assignment. Alternatively, if instances are not sensitive to color, then performance on these patterns should remain at the practiced level of performance on the original items. If a general procedure is being learned, such as that suggested by process-based theories, performance should show improvement independent of the type of transfer pattern. The degree of importance of color configuration for instance retrieval should be reflected in the extent to which performance on old/new items resembles performance on new/new items.

Method

Subjects. Six University of Illinois undergraduates served as subjects. Each subject participated in five 45-min sessions. These subjects were paid $20 for their participation.

Apparatus and stimuli. The 60 patterns (30 training patterns and 30 novel transfer patterns) were constructed in the same manner as in our previous experiments, with the exception that asterisks were used instead of letters. Half of the asterisks were red, and the other half were green (in the case of an odd number, half of the total number less one were one color, and the remaining were the other color). An additional 30 patterns were constructed for use in the transfer test, as in our previous experiments. For these patterns, the color configuration of each of the training patterns was randomly reassigned with the constraint that the color reassignment could not be identical to the original color assignment. The spatial positions were the same as those in the original 30 training patterns. Refer to Figure 2 for an example of the patterns used in this experiment. In all, subjects saw 90 different patterns (30 throughout the 16 training blocks and at transfer, the same 30 with a different color configuration, and 30 new patterns only at transfer). Patterns were displayed, and responses were collected in the same manner as in our previous experiments.

Procedure. Practice, training, and transfer trials were conducted in the same manner as they were in Experiment 2a; training occurred over four sessions (60 practice trials with numbers but no patterns to count and 480 trials of the counting task at each training session), and a single transfer session included three trial types (120 trials per type: old/old, old/new, and new/new).

In each training session, following 60 practice trials (10 at each numerosity level), each of the five training exemplars generated for each level of numerosity was presented four times in each of four 120-trial blocks, for a total of 16 presentations per item. The 60 practice trials and 480 trials took approximately 45 min to complete.

The final session included four 90-trial blocks, with one presentation of each of the 30 training, novel transfer, and old/new transfer patterns in each block.

Design. Session (1–4), level of numerosity (6–11) and type of transfer pattern (old/old, new/new, and old/new) were manipulated within subject.

Results and Discussion

Mean accuracy across the four training sessions and across the six levels of numerosity was .93. None of the means across subjects, sessions, and numerosity levels was significantly different from any other.

As in our previous experiments, across training sessions response latencies decreased with practice and increased with numerosity, and the increase in response latency with number of elements per pattern was smaller as training increased across sessions. This pattern of results is consistent with analogous ANOVAs conducted on the slopes of the function relating response latency to numerosity, which show a decrease in slope with training and an increase in slope to an earlier level of training for new patterns at transfer. Given that both sets of analyses converge, for brevity, only the analyses involving slope are reported here. A table presenting RT as a function of numerosity and practice as well as ANOVA summary tables for the analyses of RT are presented in the Appendix. Goodness of linear fit in the response latency data is assessed in three ways: the proportion of treatment variance accounted for by the linear trend; the squared correlation between mean RT and numerosity for each subject, averaged across subjects. These statistics are reported in Table 2. In general, the fit is reasonably good, and thus we focus the remainder of the analysis and discussion on slopes. Again, the slope of the response latency function decreased with training. The average slope as a function of session is presented in Figure 5.

The mean slope from Session 1 was 368 ms per item, which dropped to 179 ms in Session 2, 123 ms in Session 3, and 71 ms in Session 4. In the analysis of training slope, session was significant, $F(3, 15) = 13.20$, $p < .01$, $MS_e = 7{,}611$. Post hoc Tukey pairwise comparisons revealed significant differences between Slope 1 and Slopes 2–4, $q(4, 15) > 5.25$, $p < .01$, $MS_e = 7{,}611$, and no differences between Slopes 2, 3, and 4, $q(4, 15) < 4.08$, $p > .05$, $MS_e = 7{,}611$, which suggests that performance had reached asymptote by Session 4.

At transfer, mean accuracy across numerosity level (6–11) for each stimulus type was as follows: old/old, .93; new/new, .91; and old/new, .95. None of the means across stimulus type and numerosity level was significantly different from any other. Response latency was highest for new/new patterns, followed by that for old/new and old/old patterns, and response latency increased with numerosity. Again, response latency slopes decreased from new/new to old/new patterns and from old/new to old/old patterns. The slope for old/old transfer patterns was 72 ms per item; for old/new patterns, 118 ms per item; and for new/new patterns, 307 ms per item. The effect of stimulus type on slope was significant, $F(2, 10) = 15.98$, $p < .01$, $MS_e = 5{,}841$. Tukey post hoc contrasts revealed that slopes for old/old training patterns and old/new transfer patterns were not significantly different from each other, $q(3, 10) < 3.88$, $p > .05$, $MS_e = 5{,}841$, whereas the mean slopes for both old/old and old/new patterns were significantly different from the mean slope for new/new patterns, $q(3, 10) > 3.88$, $p < .05$, $MS_e = 5{,}841$.

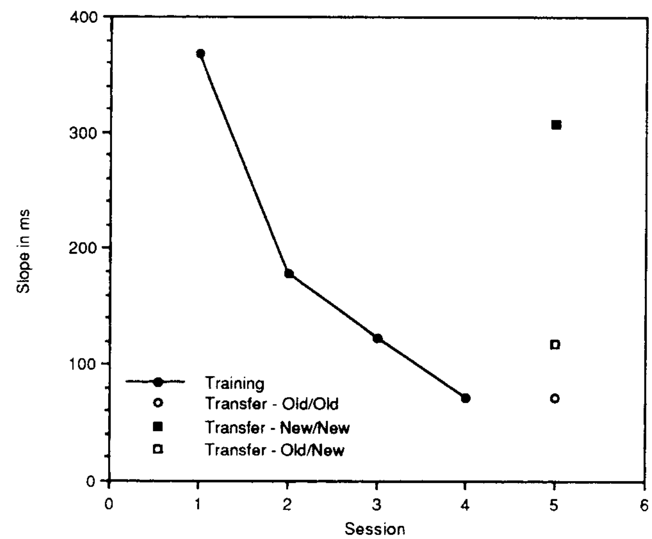

Figure 5. Slope of function relating reaction time to numerosity as a function of training session from Experiment 3a. (Sessions 1–4 used a single set of stimulus patterns for each subject; novel patterns were used to test the item specificity of learning in Session 5.)

The results of this study are consistent with results from our previous experiments, in that learning was item specific, because the improvement in performance obtained across training did not transfer to novel patterns. Also, good transfer to old/new patterns suggests that organization in terms of color is not preserved in the representation of an instance.

Experiment 3b

The results of Experiment 3a suggest that organization of a pattern in terms of color is not preserved in the representation of an instance. In this experiment, performance on old/new transfer items (patterns adhering to the same spatial arrangements as those experienced during training but having a novel color configuration) was as good as performance on old/old (training) items. Experiment 3b was conducted to replicate this finding at an intermediate stage of training in the counting task. Again, as performance becomes driven by instance retrieval rather than by algorithmic processing, the identity of stimulus elements, although unimportant in the counting algorithm, may play a more important role in instance retrieval.

The same hypothesis tested in Experiment 3a is tested here: If the representation of an instance is organized in terms of groups, with perceptual groups created by color, then performance on old patterns with a different color assignment (old/new) should be poor, because a change in the constituent structure of the grouping (changing the color assignment for old/new patterns) should impair memory retrieval. Alternatively, if color configuration is not preserved in the representation of an instance, then performance on old/new patterns will remain at the practiced level of performance on the original old/old items. If performance is driven by learning a general procedure, rather than by instance retrieval, then performance should be equally good for all types of transfer patterns. The degree of importance of color configuration for instance retrieval should be reflected in the extent to which performance on old/new items resembles performance on new/new items.

Method

Subjects. Twenty-four University of Illinois undergraduates served as subjects.[2] Each subject participated in a 1-hr session, either for pay or for credit in an introductory psychology course.

Apparatus and stimuli. The 90 patterns (30 old/old training patterns, 30 new/new novel transfer patterns, and 30 old/new patterns using spatial configuration of the old/old training patterns but with a different color configuration) were constructed in the same manner as in Experiment 3a. Patterns were displayed, and responses were collected in the same manner as in our previous experiments.

Procedure. Practice, training, and transfer trials were conducted in the same manner as in previous single-session experiments (60 practice trials with numbers only and no patterns to count, 480 counting trials, and 90 transfer trials, 30 each of old/old, new/new, and old/new patterns). With practice, training, and transfer, subjects completed 570 trials.

Design. Training block (1–16), level of numerosity (6–11), and type of transfer pattern (old/old, new/new, and old/new) were manipulated within subject.

Results and Discussion

The mean accuracy across the 16 training blocks was .94. At transfer, the mean accuracy was .94 for old patterns, .92 for novel patterns, and .94 for old patterns with a new color configuration. None of the numbers was significantly different from any other.

As in the preceding experiments, RT data were analyzed by calculating the slopes of linear regression functions relating RT to numerosity, justified by the pattern of response latencies. A table presenting RT as a function of numerosity and practice as well as ANOVA summary tables for the analyses of RT are presented in the Appendix. Goodness of linear fit in the response latency data is assessed in three ways: the square of the correlation between raw RT and numerosity for each subject, averaged across subjects; the squared correlation between mean RT and numerosity; and the proportion of treatment variance accounted for by the linear trend. These statistics are reported in Table 3.

Two sets of slopes were calculated and analyzed. The first included mean RTs at all numerosity levels (6–11). The second was calculated after dropping RTs from the endpoints, 6 and 11, and was motivated by a concern that performance in this single session may be more susceptible to strategy-use given that complete automaticity does not obtain in a single session. Results from the two training analyses converged. The original set of slopes and analysis are reported first. The average slopes across training block and at transfer are presented in Figure 6. In the initial block, the mean slope was 380 ms per item, dropping to 282 ms per item in the final block. The main effect of training block was significant, $F(15, 345) = 1.948$, $p < .05$, $MS_e = 9{,}408$. A linear contrast analysis conducted on training slopes showed a significant effect of training block (1–16), $F(1, 23) = 7.032$, $p < .05$, $MS_e = 28{,}929$, which suggests that performance did improve with practice.

At transfer, the slope was 286 ms per item for old patterns, 315 ms per item for novel (new/new) patterns, and 284 ms per item for training patterns with a new color configuration (old/new). An ANOVA was conducted on transfer slopes with type of transfer pattern (old/old, new/new, old/new) as a factor. The main effect of type of transfer pattern was not significant, $F(2, 46) < 1$, $p = .48$, $MS_e = 9{,}794$, which revealed that there were no significant differences between any of the transfer slopes. Two additional contrasts—between initial slope (Block 1) and slope for new patterns at transfer (380 ms per item and 315 ms per item, respectively) and between final training slope (Block 16) and slope for old patterns at transfer (282 ms per item and 286 ms per item, respectively)—were not significant.

Slopes calculated using the second set of slopes, calculated from the 7, 8, 9, and 10 numerosity levels, ranged from 456 ms per item in the initial training block to 237 ms per item by the final training block. The main effect of training block approached significance, $F(15, 285) = 1.620$, $p = .07$, $MS_e = 28{,}427$. A linear contrast analysis showed a significant effect of training block, $F(1, 19) = 6.175$, $p < .05$, $MS_e = 51{,}707$. At transfer, the latency slopes were 199 ms per item (old/old), 366 ms per item (new/new), and 215 ms per item (old/new). An ANOVA was conducted on transfer slopes with type of transfer pattern as a factor. In contrast to the first analysis (using slopes calculated from numerosity levels 6–11), there was a significant effect of type of transfer pattern, $F(2, 38) = 5.931$, $p < .01$, $MS_e = 28{,}865$. Tukey post hoc contrasts revealed a significant difference between the new patterns (new/new) and the old patterns (old/old and old/new), $q(3, 38) > 3.49$, $p < .05$, $MS_e = 28{,}865$, but no statistical difference between the two types of old patterns. The same two additional contrasts presented for the analyses of the original set of slopes—between initial slope (Block 1) and slope for new patterns at transfer (456 ms per item and 366 ms per item, respectively) and between final training slope (Block 16) and slope for old patterns at transfer (237

[2] We did not test subjects in Experiment 3b for color blindness, but note that only some 8% of the male population are color blind (Mollon, 1982) and that instructions explicitly mentioned that subjects would be counting red and green dots.

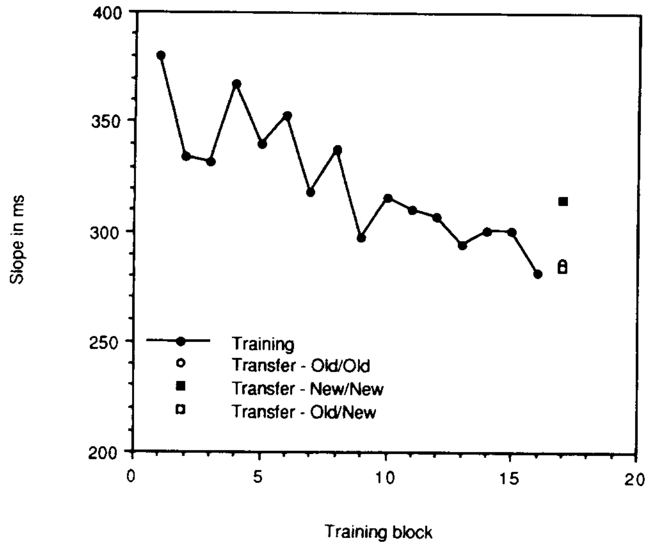

Figure 6. Slope of function relating reaction time to numerosity as a function of training block (1–17) from Experiment 3b. (New/new and old/new items were included in Block 17 for a transfer test.)

ms per item and 199 ms per item, respectively)—were not significant.

These results suggest that learning was item specific, because the improvement in performance did not transfer to novel patterns. Performance on these novel patterns was not statistically different from initial performance, although the difference in mean slope between the initial training block and new patterns at transfer was perhaps too large to ignore (90 ms per item). Good transfer to old/new patterns suggests that organization in terms of color is not preserved in the representation of an instance; taken with the results of Experiments 2a, 2b, and 3a, so far only spatial location of the individual elements composing each pattern seems to be maintained in memory, but not identity or color. Experiments 4a and 4b addressed a third aspect of representation, the organization of the pattern as a whole. Changing the orientation of training patterns at transfer changes the organization of a pattern as a whole but maintains the relations among the component parts (Rock, 1973).

Experiment 4a

This experiment is motivated by the idea that the representations of the patterns are held as whole units in visual–spatial memory. This idea is tested by manipulating orientation, which is known to affect visual–spatial memory from the seminal work of Rock (1973). He found that changing the orientation of familiar visual patterns impaired performance in tests of recognition memory. Experiment 4a assesses the necessity of preserving the spatial orientation of a pattern for supporting memory retrieval of a trained instance, in the same manner as the preservation of color configuration was assessed in Experiments 2a, 2b, 3a, and 3b. Training was conducted as in previous multiple-session experiments, and both trained (old/old) and new (new/new) items were presented for a transfer test. In addition, the third type of transfer pattern (old/new) was created by rotating each trained pattern 180° about its center, thereby preserving the spatial relations among the elements but changing the orientation of the pattern within the reference frame of the display. Examples of the three types of patterns are presented in Figure 2. If instances are held in visual–spatial memory and thus are subject to manipulations that affect visual–spatial memory, then rotated (old/new) patterns should produce performance comparable to that of novel patterns. Old/new performance falling between old/old and new/new performance indicates that orientation has some importance for instance retrieval.

Method

Subjects. Six University of Illinois undergraduates served as subjects. Each subject participated in five 45-min sessions. These subjects were paid $20 for their participation.

Apparatus and stimuli. The 90 patterns formed of white asterisks (30 old/old training patterns, 30 new/new novel transfer patterns, and 30 old/new patterns formed by rotating each old/old training pattern 180° about its center) were constructed in the same manner as in the previous experiments. Patterns were displayed, and responses were collected in the same manner as they were in our previous experiments.

Procedure. Practice, training, and transfer trials were conducted in the same manner as in Experiments 2a and 3a. Training occurred over four sessions (60 practice trials with numbers but no patterns to count, 480 trials of the counting task at each training session), and one transfer session included three trial types (120 trials per type: old/old, old/new, and new/new).

In each training session, after 60 practice trials (10 at each numerosity level), each of the five training exemplars generated for each level of numerosity was presented four times in each of four 120-trial blocks, for a total of 16 presentations per item. The 60 practice trials and 480 trials took approximately 45 min to complete.

The final session included four 90-trial blocks, with one presentation of each of the 30 training, novel transfer, and old/new transfer patterns in each block.

Design. Session (1–4), level of numerosity (6–11), and type of transfer pattern (old/old, new/new, and old/new) were manipulated within subject.

Results and Discussion

Mean accuracy across training session and numerosity was .94. None of the means across subjects, sessions, and numerosity levels was statistically different from any other.

The same pattern of results obtained in earlier experiments was exhibited in this experiment: (a) an increase in response latency with an increase in numerosity; (b) a decrease in response latency with an increase in training; and (c) a smaller increase in response latency with numerosity as training progressed. A table presenting RT time as a function of numerosity and practice as well as ANOVA summary tables for the analyses of RT are presented in the Appendix. Goodness of linear fit in the response latency data is assessed in three ways: the proportion of treatment variance accounted for by the linear trend, the squared correlation between mean RT and numerosity, and the square of the correlation between raw RT and numerosity for each subject, averaged across subjects. These statistics are reported in Table 2.

As in previous multisession experiments, slope decreased with session, from 378 ms per item in Session 1 to 279 ms per item in Session 2, 182 ms per item in Session 3, and 106 ms per item in Session 4. The average slope as a function of session is presented in Figure 7. The effect of session on training slope was significant, $F(3, 15) = 11.13, p < .01, MS_e = 7{,}521$. Tukey HSD post hoc contrasts revealed significant differences between Slopes 1 and 3, Slopes 1 and 4, and Slopes 2 and 4, $q(4, 15) > 5.25, p < .01, MS_e = 7{,}521$, which suggests that the decrease in slope across sessions had reached asymptote by Session 3. Pairs of successive slopes were not significantly different from each other.

At transfer, there was a significant effect of stimulus type on accuracy, unlike in previous experiments, $F(2, 10) = 7.54$, $p < .05$, $MS_e = 21.48$, such that accuracy was greatest for old/old patterns (.96), followed by old/new patterns (.94), and then new/new patterns (.91). Neither numerosity level nor the interaction between numerosity and stimulus type was significant. The pattern of accuracies across stimulus type is consistent with earlier response latency results, in which latency is highest for new/new, followed by old/new, and then by old/old patterns. In this experiment, the effect of stimulus type on response latency was only marginally significant, $F(2, 10) = 4.18, p < .10, MS_e = 799{,}683$, which suggests that accuracy, rather than response latency, was affected by stimulus type. Rather than slowing down on new/new patterns, subjects suffered a decrease in accuracy. The effect of numerosity on response latency was again significant, $F(5, 25) = 15.08, p < .01, MS_e = 349{,}764$, such that response latency increased with numerosity. The interaction between stimulus type and numerosity was not significant.

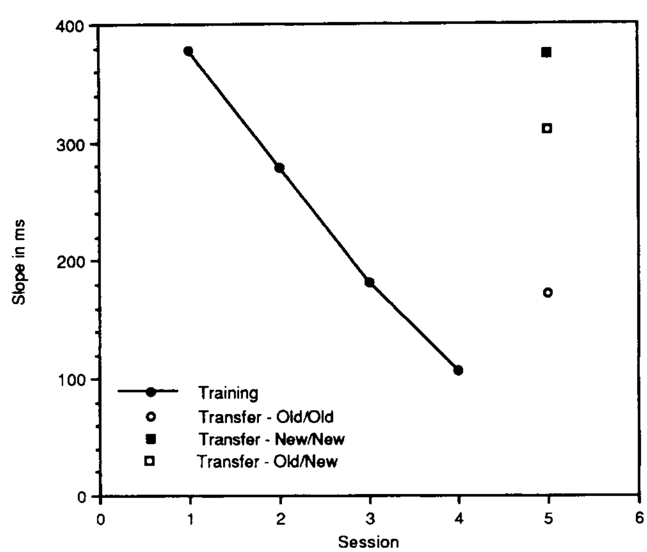

Figure 7. Slope of function relating reaction time to numerosity as a function of training session from Experiment 4a. (Sessions 1–4 used a single set of stimulus patterns for each subject; novel patterns were used to test the item specificity of learning in Session 5.)

The slope of the response latency function for old/old patterns was 172 ms per item; for old/new patterns, 310 ms per item; and for new/new patterns, 375 ms per item. There was a marginally significant effect of stimulus type on slope, $F(2, 10) = 3.65, p < .10, MS_e = 17{,}641$. Post hoc contrasts revealed that slopes for the two types of new patterns, old/new and new/new, were not significantly different from each other, $q(3, 10) < 3.88, p > .05, MS_e = 17{,}641$. The difference between the mean slopes for both types of new patterns and old/old patterns was significant, $q(3, 10) > 3.88, p < .05$, $MS_e = 17{,}641$.

The results of this study are consistent with results from our previous experiments, in that learning was item specific, because the improvement in performance obtained across training did not transfer to novel patterns. When we changed the orientation of learned patterns, performance was impaired, which suggests that orientation of each pattern as a whole is important in memory.

Experiment 4b

The results of Experiment 4a suggest that the orientation of each pattern as a whole is preserved in the representation of an instance, because changing the orientation of learned patterns impaired performance. Experiment 4b was conducted to replicate this finding at an intermediate stage of training in the counting task. This experiment used the same practice, training, and transfer procedure used in Experiments 2b and 3b. Again, as in these experiments, transfer was tested in a final session, but the same three types of stimulus patterns used in Experiment 4a were presented: patterns used during training (old/old); novel patterns seen only at transfer (new/new); and patterns constructed by rotating each old/old training patterns 180° about its center (old/new). The same hypothesis tested in Experiment 4a is tested here, but at an intermediate stage of training: If instances are held in visual–spatial memory, and thus are subject to manipulations that affect visual–spatial memory, then old/new rotated patterns should produce performance comparable to that of novel patterns. Old/new performance falling between old/old and new/new performance indicates that orientation has some importance for instance retrieval; old/new performance no different from old/old performance suggests that orientation plays no role in instance retrieval.

Method

Subjects. Twenty-four University of Illinois undergraduates served as subjects, participating either for pay or for credit in an introductory psychology course. Participation involved a single session that lasted approximately 1 hr. The data from 2 subjects were not included in any analyses, because these subjects performed at an error rate above the predetermined cutoff of 15%. These subjects were not replaced.

Apparatus and stimuli. The 30 old/old and 30 new/new patterns were constructed in the same manner as in Experiment 4a, such that all patterns were composed of white asterisks. Thirty old/new patterns were formed by rotating each old/old training pattern

180° about its center. Patterns were displayed, and responses were collected in the same manner as described for previous experiments.

Procedure. Practice, training, and transfer tests were all conducted in the same manner as in the previous single-session experiments. There were three types of transfer patterns and thus a total of 570 trials (60 practice trials with numbers but no patterns to count, 480 training trials, and 90 transfer trials with dot patterns).

Results and Discussion

The mean accuracy across the 16 training blocks was .95. Accuracy for the three types of transfer patterns was as follows: old/old, .92; new/new, .92; and old/new, .92. The same pattern of response latency results described for previous experiments obtained in this experiment, justifying an analysis of the slopes of linear regression functions relating RT to numerosity. A table presenting RT as a function of numerosity and practice as well as ANOVA summary tables for the analyses of RT are presented in the Appendix. Goodness of linear fit in the response latency data is assessed in three ways: the proportion of treatment variance accounted for by the linear trend; the squared correlation between mean RT and numerosity; and the square of the correlation between raw RT and numerosity for each subject, averaged across subjects. These statistics are reported in Table 3.

Two sets of slopes were computed and analyzed as in the previous single-session experiments, but in this case, there was a discrepancy between the two analyses regarding the transfer results. It is reasonable to believe that the discrepancy involves the use of the slopes calculated only from numerosity levels 7–10, as explained later. The analyses using the first set of slopes, calculated using all six levels of numerosity, are presented first.

The average slopes calculated using all numerosity levels, 6–11, across training block and at transfer, are presented in Figure 8. The initial slope was 349 ms per item, which dropped to 272 ms per item by the final training block. The main effect of training block (1–16) was significant, $F(15, 315) = 3.108$, $p < .001$, $MS_e = 9{,}651$, and a linear contrast reflecting the reduction in slope over blocks was significant, $F(1, 21) = 8.704$, $p < .01$, $MS_e = 39{,}409$, which suggests that performance did improve with practice.

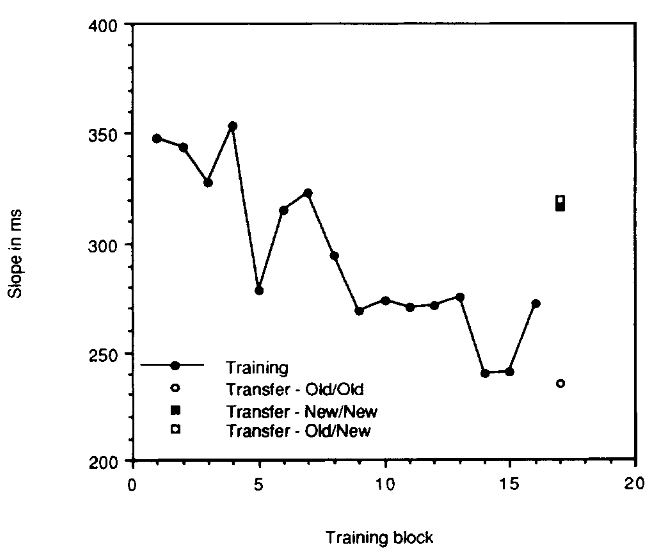

Figure 8. Slope of function relating reaction time to numerosity as a function of training block (1–17) from Experiment 4b. (New/new and old/new items were included in Block 17 for a transfer test.)

At transfer, the slope was 235 ms per item for old patterns, 317 ms per item for novel (new/new) patterns, and 320 ms per item for training patterns with a 180° change in orientation (old/new). An ANOVA was conducted on transfer slopes with type of transfer pattern (old/old, new/new, old/new) as a factor. The main effect of type of transfer pattern was significant, $F(2, 42) = 5.195$, $p < .01$, $MS_e = 9{,}873$. Post hoc contrasts revealed a significant difference between the old patterns (old/old) and the new patterns (new/new and old/new), $q(3, 42) > 3.44$, $p < .05$, $MS_e = 9{,}873$, but no statistical difference between the two types of new patterns, $q(3, 42) < 3.44$, $p > .05$, $MS_e = 9{,}873$. Two additional contrasts—between initial slope (Block 1) and slope for new patterns at transfer (349 ms per item and 317 ms per item, respectively) and between final training slope (Block 16) and slope for old patterns at transfer (272 ms per item and 235 ms per item, respectively)—were not significant.

The results of the analysis of the training data using slopes calculated from the intermediate numerosity levels were consistent with the initial analysis of training data using slopes calculated from all numerosity levels. Slopes ranged from 438 to 270 ms per item. The main effect of training block approached significance, $F(15, 315) = 1.537$, $p = .09$, $MS_e = 40{,}239$, and there was a significant linear trend over training block, $F(1, 21) = 4.915$, $p < .05$, $MS_e = 45{,}938.392$. The discrepancy involved the transfer data. Transfer slopes calculated from the intermediate numerosity levels (7–10) were as follows: 277 ms per item for old/old patterns; 275 ms per item for new/new patterns; and 352 ms per item for old/new patterns. Comparisons between types of transfer items showed no difference between the old/old items and the new/new items, $F(1, 21) = 0.001$, $MS_e = 1{,}282{,}076$, nor between the old/new items and the other two types of items, $F(1, 21) = 2.772$, $MS_e = 183{,}272$, although this difference approached significance, $p = .11$.

It was rather surprising that in this case the slope for novel patterns was not only lower than that for training items that had been rotated but was also lower than that for the original training items that had each been presented 16 times in training. The validity of using slopes rather than actual RTs was checked in this case, as in other experiments, by analyzing the RT data for a linear trend over numerosity level. Whereas in all other cases the linear trend accounted for a significant amount of the variance in RTs, which justifies the use of slopes as a unit of analysis, in this case there were significant departures from linearity, $F(4, 84) = 15.83$, $p < .05$, $MS_e = 219{,}597.39$. This suggests that any change in slope may not reflect a comparable change in RT. A 4 (level of numerosity, 7–10) × 3 (type of transfer pattern) ANOVA was conducted on the transfer RTs. There was a significant main effect of type of transfer pattern, $F(1, 21) = 6.946$, $p < .05$, $MS_e = 65{,}084.961$. Post hoc contrasts revealed that response times to new/new patterns were not significantly different from those to new/old patterns, $F(1, 21) = 0.488$, $p = .493$, $MS_e = 38{,}955.139$, and response times to old/old patterns were significantly less than those to both new/old and new/new patterns, $F(1, 21) = 8.493$, $p < .01$, $MS_e = 518{,}791.133$. This pattern of results corresponds to that evidenced by the first analyses, using slopes computed from all numerosity levels, and is also consistent with the results of Experiment 4a. The slope for completely novel patterns is significantly higher than that for training patterns. The elevated slope for new/new patterns in this second analysis (using only levels 7–10) may be considered anomalous (brought about perhaps by an elevated response latency to new/new patterns composed of seven elements), and the results of the first analyses are used in drawing conclusions.

As did the results of all of our previous experiments, these results suggest that learning was item specific, because the improvement in performance did not transfer to novel patterns. Performance on these novel patterns was not statistically different from initial performance. In addition, when we changed the orientation of learned patterns, performance was impaired substantially, which suggests that the orientation of the pattern as a whole was preserved in the memory representation.

General Discussion

This article began with two questions. First, can memory-based theories account for learning and transfer of automaticity with nonsymbolic stimuli? Second, what is the nature of the memory representation that underlies automatization? In this section, we discuss the answers that the experiments provide to each of these questions.

Memory-Based Account of Automaticity With Nonsymbolic Stimuli

Memory-based models of automaticity predict narrow transfer, because in this case, learning involves the storage of specific instances in memory, which, with practice, are directly retrieved rather than computed. The data from all seven of our experiments suggest item-specific learning, which is predicted by memory-based theories, because in all cases transfer performance on unstudied items, which were constructed in the same manner as learned patterns, was significantly slower than performance on the learned items.

In Experiment 1, performance reached asymptote by the fourth session, after significant improvements in performance over the first two sessions; and at transfer, performance regressed to the level displayed during the second session. These results suggest that automatic performance in the counting task had indeed been obtained and that this performance was instance-based. Experiment 1 demonstrated that the 50–100-ms-per-item reductions in slope seen in Experiments 2b, 3b, and 4b represent steps on the way to elimination of slope with further practice. In Experiments 2a, 3a, and 4a, slope was substantially reduced (109, 72, and 106 ms per item, respectively). The same pattern of results was observed with intermediate and extended practice. This is consistent with the instance theory idea that automaticity effects depend on what is in memory, not on how much practice subjects have had. Apparently, subjects had acquired some representation of the stimuli in the single-session experiments, and that representation was not substantially different from what developed after four sessions of practice.

Nature of the Memory Representation Underlying Automatization

If the memory-based view is to gain acceptance as a theory of automaticity, a clearer understanding of the memory representation needs to be obtained. Representation should be an important issue for any theory of automaticity. Its importance is brought to the fore in memory-based theories.

It seems that the identity of individual elements within a stimulus pattern is not important in the memory representation, because there was transfer to old patterns with new elements in Experiments 2a and 2b. The spatial location of each element within the pattern is salient in the representation, though, because presenting patterns with a novel spatial configuration that nonetheless maintained the identity of elements from the training patterns negatively influenced performance (Experiments 2a and 2b).

From Experiments 3a and 3b, we know that color is not important, because transfer performance on patterns that maintained the spatial organization of a pattern experienced during training, but displayed a novel color configuration, was almost as good as performance on the original patterns with the original color configuration. In comparing performance on the three types of patterns (old/old, old/new, and new/new), we find that the decrement in performance on new/new but not old/new patterns is consistent with the results of Experiments 2a and 2b. This suggests that spatial organization is represented in memory and, furthermore, that color is not represented. These findings should be interpreted with caution, though, because they are based on null results. Also, although our experiments provide no evidence that these particular manipulations of identity and subgrouping are important, it is possible that stronger manipulations would have an effect (e.g., switching from small dots to large ones or switching from white-on-black to black-on-white to manipulate element identity; using displays in which colors lead subjects to arrange elements first by columns and then switching colors so that subjects are led to arrange elements by rows to manipulate subgrouping).

The lack of transfer to patterns that contained the same spatial relations among items but were rotated 180° in Experiments 4a and 4b suggests that the configuration of the pattern as a whole, in addition to the spatial location of individual elements in relation to each other, is important to the memory representation, because changing the orientation of the pattern as a whole interfered with memory retrieval. Again, as in our previous experiments, there was poor transfer to completely new patterns, which suggests that learning that occurred with practice was for the most part specific to items experienced during training.

Across these experiments, novel patterns at transfer produced, in general, shallower slopes than those observed in the first block or first session. This positive transfer of learning to novel patterns can be accounted for in at least three ways. First, subjects may have simply learned to count faster. Second, new patterns may have been similar to training patterns on some attribute that caused training instances to be retrieved and used in making the numerosity judgment. The use of training instances during performance with novel patterns may have reduced RTs and slopes. Third, subjects may have used memory associations between elements of a pattern and digits on the number line to speed up their selection of the next item to count. Such associations may have then built a "path" through the pattern that became progressively easier as practice continued. These alternatives are suggested to emphasize that we do not claim that all learning is instance-based, but rather that much of it is. Although instance learning may not be the single mechanism driving performance, it is clearly important (see Logan, 1988a).

In addition to positive transfer of learning to novel patterns, transfer was less than complete with changes in element identity and color and not completely absent with changes in orientation. The incomplete transfer with identity or color changes could be driven by identity or color information contained by certain training patterns. Identity or color information may have made a pattern more memorable. For example, a pattern constructed with the letter *o* as elements could have been shaped like that letter, or color may have been assigned to elements within a pattern such that all of the green asterisks were on the right side of the screen and all of the red asterisks were on the left. Changing the identity or color of these patterns may have affected transfer, although it may not have for other patterns. Similarly, transfer even with a change in orientation may have resulted for patterns that could be recognized independently of orientation (e.g., patterns that are horizontally and vertically symmetric). Alternatively, a pattern may have been encoded independent of orientation, perhaps by focusing on pairwise relations between elements.

Across these experiments, it becomes clear that the representation of an instance does not preserve all of the information present in the stimulus, but rather seems to be constrained by the nature of the task and by an attentional filter. Attributes important to the task are likely to be attended to and thus incorporated into the memory trace; less important attributes may be ignored and thus not incorporated into the trace. Instance theory assumes that encoding is an obligatory consequence of attention—that is, attention to an item or event causes it to be encoded into memory. The algorithm responsible for initial performance produces memory traces that eventually come to support performance without the algorithm, and it is the nature of the specific algorithm that determines to what the subjects attend. In this case, the counting algorithm involves three components: spatial indexing, mapping from the indexed element to an internal representation of the number line, and producing a response. The most likely candidates for focus of the attentional mechanism include the spatial position of each element in the display, its corresponding value on the number line, and the final value. Each of these corresponds to a component of the algorithm. Neither the identity nor the color of the individual elements is in any way relevant to spatial position or numerical value (Experiments 2a and 2b, Experiments 3a and 3b), but a change in orientation of an entire pattern would indeed affect the first component, spatial indexing (Experiments 4a and 4b).

In the second component of the counting task, mapping each indexed element to an internal representation of the number line, subjects associate particular digits with particular elements in the display. When no elements remain to be indexed and mapped, the final digit counted is associated with the pattern as a whole. The former associations, element-to-digit, are not useful beyond leading to the next step of the algorithm, because element identity cannot predict numerosity. Element-to-digit associations, therefore, cannot replace the counting algorithm in producing a numerosity response. The latter pattern-to-digit associations, in contrast, can replace the counting algorithm by producing the same result, the numerosity of a pattern. Element-to-digit associations can, however, assist in the previous component of the counting task, spatial indexing. When an element is indexed and mapped to a digit on the number line, that digit may serve to prime the next digit to be mapped. The next digit, in turn, may serve to prime the element in the pattern that had in the past been mapped to it as the next element in the display to be indexed. In this manner, the counting algorithm could speed up because spatial indexing speeds up. Configurations of elements within a pattern can assist spatial indexing in the same manner as element-to-digit associations. In this sense, more than just the numerosity of a pattern is remembered, and thus something other than retrieving numerosity can speed up performance. Configurations of elements may be common even between training and novel transfer items, which suggests yet another way that performance can be driven by something other than retrieval of numerosity in its strictest interpretation (remembering a numerosity response; Barsalou, 1990).

Instance theory may not be the only memory-based account of the data presented here. These data are consistent with memory-based theories that assume that the representation includes configurations of properties of the stimuli. Instance theories are clearly of this type, because they record the particular combination of properties to which subjects attend. Strength theories could be of this type, in principle, depending on how they are construed. Perhaps some sort of memory-assisted processing is responsible for performance in the initial session. Across Experiments 2b, 3b, and 4b, the pattern of initial learning is perhaps better characterized by some intermediate mechanism by which memory guides the visual scan path through the display. In this case, the pattern acts as a retrieval cue, not for the solution, but for the most efficient visual scan path through the display (Noton & Stark, 1971). The retrieved path is item specific, which distinguishes this alternative from a purely process-based account. This would be indicative of an amalgam of process- and memory-based improvement, because memory retrieval is supporting the counting process by producing a learned visual path by which to index the array. Instead of a single transition from algorithmic computation to memory retrieval, the memory-assisted algorithm view suggests two transitions: one from the algorithm to memory-assisted algorithm and another, which occurs later in training, from memory-assisted algorithm to instance retrieval. First, subjects initially rely on an algorithm, and then by the end of the initial session, they use memory to help them execute the algorithm. Eventually, as performance becomes fully automatized, the algorithm is abandoned, and subjects rely on memory entirely. The notion of memory-assisted processing provides an interesting alternative to the original idea proposed by instance theory of two mechanisms, an algorithm and memory retrieval. A series of algorithms ultimately ending in memory retrieval is consistent with the broad conception of instance theory, though it may make the formal analysis much more difficult. More work is necessary to specifically test this possibility. Regardless of whether instance theory provides the only account, though, any account one offers must deal with the phenomenon of item-specific learning. If memory assists the counting algorithm, it assists only for patterns experienced during training. Memory-assisted algorithmic processing does not transfer to new patterns.

References

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

Barsalou, L. W. (1990). On the indistinguishability of exemplar memory and abstraction in category representation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 3, pp. 61–88). Hillsdale, NJ: Erlbaum.

Boring, E. G. (1950). *A history of experimental psychology* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Chi, M. T. H., & Klahr, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology, 19,* 434–439.

Jensen, E. M., Reese, E. P., & Reese, T. W. (1950). The subitizing and counting of visually presented fields of dots. *The Journal of Psychology, 30,* 363–392.

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. Pomerantz (Eds.). *Perceptual organization* (pp. 181–211). Hillsdale, NJ: Erlbaum.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology, 62,* 498–525.

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

Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. Special Issue: Skill. *Canadian Journal of Psychology, 39,* 367–386.

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

Logan, G. D. (1988b). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors, 30,* 583–598.

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

Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 883–914.

Logan, G. D. & Klapp, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 179–195.

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General, 111,* 1–22.

Mollon, J. D. (1982). Colour vision and colour blindness. In H. B. Barlow & J. D. Mollon (Ed.), *The senses: Cambridge texts in the physiological sciences* (Vol. 3, pp. 165–191). Cambridge, England: Cambridge University Press.

Neter, J., Wasserman, W., & Kutner, M. H. (1985). *Applied linear statistical models: Regression, analysis of variance, and experimental design* (2nd ed.). Homewood, IL: Irwin.

Noton, D., & Stark, L. (1971). Eye movements and visual perception. *Scientific American, 224,* 34–43.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial index model. *Cognition, 32,* 65–97.

Rock, I. (1973). *Orientation and form.* San Diego, CA: Academic Press.

Schneider, W., & Fisk, A. D. (1984). Automatic category search and its transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 1–15.

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

Shiffrin, R. M., & Dumais, S. T. (1981). The development of automatism. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 111–140). Hillsdale, NJ: Erlbaum.

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

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

(Appendix begins on next page)

Appendix

Reaction Time Analyses

For all experiments, two ANOVAs were conducted on response latencies: The first examined training response latencies for an effect of numerosity (6–11) and practice (Session 1–12 or 1–4 in the multisession experiments; Training Block 1–16 in the single-session experiments); the second examined transfer response latencies for an effect of numerosity (6–11) and stimulus type (Sessions 1–12 vs. Session 13 for Experiment 1; old/old, old/new, and new/new for all other experiments). Table A1 presents the mean response latencies, and Table A2 presents the summaries of these ANOVAs. Across experiments, the same pattern of results emerged: (a) Response latency increased with numerosity; (b) response latency decreased with training; (c) the increase in response latency with numerosity was smaller as training progressed; (d) response latency increased with numerosity at transfer as well as training; (e) response latency was highest for new (or new/new) transfer patterns, followed by old/new patterns in Experiments 2–4, with old (or old/old) patterns exhibiting the lowest response latency; and (f) the increase in response latency with numerosity was smaller for old patterns than for new patterns. The one exception, a main effect of transfer stimulus type on accuracy but not response latency, occurred in Experiment 4a.

Table A1
Mean Response Latencies (in Milliseconds) as a Function of Numerosity and Training in Experiments 1, 2a, 2b, 3a, 3b, 4a, and 4b

	Numerosity level					
Training level	6	7	8	9	10	11
			Experiment 1			
Session 1	1,419	1,885	2,400	2,833	2,989	2,903
Session 2	1,186	1,486	1,803	2,289	2,360	1,693
Session 3	1,020	1,156	1,456	1,754	1,720	1,238
Session 4	949	998	1,172	1,457	1,234	983
Session 5	881	944	1,060	1,232	1,128	890
Session 6	842	931	981	1,161	1,043	906
Session 7	794	912	953	1,107	1,009	878
Session 8	765	840	894	978	890	787
Session 9	751	826	860	969	902	800
Session 10	751	776	813	937	872	776
Session 11	718	820	835	939	904	766
Session 12	702	749	788	880	880	748
Session 13 (Transfer)	1,236	1,503	1,823	2,204	2,360	2,316
			Experiment 2a			
Session 1	1,471	2,029	2,452	3,030	3,111	3,304
Session 2	1,205	1,705	2,023	2,474	2,531	2,528
Session 3	1,041	1,471	1,643	2,023	2,006	2,001
Session 4	931	1,317	1,478	1,729	1,595	1,476
Session 5: Old/old	935	1,406	1,640	1,761	1,793	1,781
Session 5: New/new	1,423	1,911	2,557	2,776	3,336	3,153
Session 5: Old/new	1,013	1,441	1,667	1,870	1,938	2,144
			Experiment 2b			
Block 1	2,032	2,513	2,964	3,654	3,365	3,589
Block 2	1,585	2,190	2,739	3,198	3,179	3,538
Block 3	1,694	2,257	2,687	3,087	3,294	3,509
Block 4	1,507	2,173	2,862	3,156	3,416	3,535
Block 5	1,468	1,972	2,675	2,912	3,200	3,198
Block 6	1,393	2,086	2,510	2,916	3,095	3,165
Block 7	1,462	1,920	2,553	2,769	3,034	3,090
Block 8	1,392	1,889	2,516	2,689	2,805	3,008
Block 9	1,407	1,909	2,365	2,536	2,958	2,977
Block 10	1,346	1,885	2,414	2,786	2,869	2,976
Block 11	1,339	1,895	2,564	2,937	2,796	2,953
Block 12	1,323	1,862	2,374	2,822	2,777	2,847
Block 13	1,299	1,817	2,370	2,682	2,730	2,766
Block 14	1,282	1,850	2,251	2,640	2,636	2,837
Block 15	1,439	1,765	2,297	2,548	2,617	2,475
Block 16	1,290	1,818	2,320	2,569	2,655	2,647

Table A1 *Continued*

Training level	Numerosity level 6	7	8	9	10	11
Transfer: Old/old	1,393	1,856	2,334	2,643	2,573	2,733
Transfer: New/new	1,711	2,170	2,784	2,967	3,275	3,410
Transfer: Old/new	1,357	1,916	2,399	2,709	2,724	2,801
		Experiment 3a				
Session 1	1,532	2,024	2,590	2,907	3,236	3,319
Session 2	1,220	1,536	1,971	1,909	2,099	2,153
Session 3	1,015	1,358	1,490	1,424	1,664	1,714
Session 4	841	1,114	1,178	1,082	1,151	1,335
Session 5: Old/old	909	1,234	1,260	1,110	1,229	1,475
Session 5: New/new	1,555	2,111	2,604	3,094	3,044	3,091
Session 5: Old/new	963	1,324	1,424	1,302	1,508	1,703
		Experiment 3b				
Block 1	1,983	2,478	3,025	3,822	3,667	3,805
Block 2	1,767	2,301	2,935	3,195	3,312	3,339
Block 3	1,654	2,374	2,906	3,074	3,373	3,200
Block 4	1,559	2,140	2,739	2,901	3,121	3,420
Block 5	1,567	2,076	2,617	2,796	3,032	3,195
Block 6	1,462	2,026	2,677	2,737	3,051	3,077
Block 7	1,450	2,033	2,442	2,614	2,976	2,958
Block 8	1,392	2,005	2,390	2,645	2,653	3,150
Block 9	1,496	2,077	2,199	2,611	2,655	3,027
Block 10	1,356	1,961	2,231	2,526	2,801	2,841
Block 11	1,292	1,891	2,238	2,373	2,739	2,818
Block 12	1,337	1,811	2,305	2,483	2,679	2,792
Block 13	1,312	1,854	2,374	2,497	2,729	2,812
Block 14	1,352	1,935	2,153	2,378	2,727	2,750
Block 15	1,287	1,897	2,186	2,395	2,703	2,707
Block 16	1,295	1,874	2,294	2,365	2,529	2,653
Transfer: Old/old	1,313	1,898	2,054	2,335	2,536	2,818
Transfer: New/new	1,564	2,160	2,376	2,697	3,023	3,057
Transfer: Old/new	1,308	1,609	2,184	2,280	2,277	2,636
		Experiment 4a				
Session 1	1,537	2,234	2,534	3,026	3,220	3,485
Session 2	1,201	1,867	1,873	2,487	2,607	2,587
Session 3	1,041	1,511	1,722	1,882	2,071	1,949
Session 4	1,165	1,407	1,632	1,836	1,736	1,681
Session 5: Old/old	1,664	1,745	1,657	1,840	2,267	2,532
Session 5: New/new	1,493	2,048	2,387	2,913	3,032	3,437
Session 5: Old/new	1,441	1,954	2,135	2,663	2,872	2,975
		Experiment 4b				
Block 1	2,032	2,610	3,189	3,344	3,705	3,782
Block 2	1,758	2,228	2,871	3,297	3,377	3,400
Block 3	1,738	2,201	2,619	3,012	3,193	3,349
Block 4	1,640	2,117	2,605	3,181	3,305	3,281
Block 5	1,682	2,097	2,622	2,835	2,952	3,062
Block 6	1,606	2,145	2,574	2,942	3,014	3,233
Block 7	1,617	1,954	2,527	2,707	3,065	3,180
Block 8	1,553	1,994	2,491	2,927	3,029	2,901
Block 9	1,678	1,927	2,361	2,738	2,937	2,861
Block 10	1,660	2,017	2,474	2,634	3,162	2,890
Block 11	1,530	1,910	2,461	2,694	2,863	2,811
Block 12	1,505	1,938	2,453	2,747	2,879	2,766
Block 13	1,510	1,859	2,395	2,711	2,888	2,772
Block 14	1,597	1,975	2,293	2,730	2,649	2,815
Block 15	1,533	1,943	2,233	2,583	2,730	2,674
Block 16	1,448	1,908	2,214	2,663	2,757	2,747
Transfer: Old/old	1,475	1,937	2,316	2,732	2,748	2,518
Transfer: New/new	1,557	2,174	2,595	2,955	2,989	3,236
Transfer: Old/new	1,567	1,947	2,646	2,936	3,026	3,104

(*Appendix continues on next page*)

Table A2
Summaries of Analyses of Variance Analyzing Response Latency as a Function of Numerosity and Training for Experiments 1, 2a, 2b, 3a, 3b, 4a, and 4b

Effect	*df*	MS_e	*F*	*p*
Experiment 1				
Training				
Session (S)	11, 33	184,342	31.50	<.001
Numerosity (N)	5, 15	152,538	10.31	<.001
S × N	55, 165	12,473	12.31	<.001
N: Linear trend	1, 3	257,249	12.80	<.050
Transfer				
Stimulus type (ST)	1, 3	346,440	43.63	<.010
Numerosity (N)	5, 15	17,330	30.04	<.001
ST × N	5, 15	17,579	18.47	<.001
N: Linear trend	1, 3	46,882	47.93	<.010
Experiment 2a				
Training				
Session (S)	3, 15	94,631	93.75	<.001
Numerosity (N)	5, 25	120,759	44.92	<.001
S × N	15, 75	33,331	8.01	<.001
N: Linear trend	1, 5	155,075	48.67	<.001
Transfer				
Stimulus type (ST)	2, 10	345,945	29.16	<.001
Numerosity (N)	5, 25	79,364	53.37	<.001
ST × N	10, 50	88,228	3.83	<.001
N: Linear trend	1, 5	75,169	81.13	<.001
Experiment 2b				
Training				
Block (B)	15, 345	319,185	25.20	<.001
Numerosity (N)	5, 115	1,001,158	155.81	<.001
B × N	75, 1,725	147,535	1.86	<.001
N: Linear trend	1, 23	2,485,257	284.25	<.001
Transfer				
Stimulus type (ST)	2, 46	317,560	28.79	<.001
Numerosity (N)	5, 115	292,470	83.74	<.001
S × N	10, 230	194,692	1.21	.286
N: Linear trend	1, 23	331,484	333.59	<.001
Experiment 3a				
Training				
Session (S)	3, 15	173,658	84.49	<.001
Numerosity (N)	5, 25	217,601	14.41	<.001
S × N	15, 75	51,277	7.66	<.001
N: Linear trend	1, 5	742,048	16.65	<.001
Transfer				
Stimulus type (ST)	2, 10	282,376	72.34	<.001
Numerosity (N)	5, 25	130,220	15.35	<.001
ST × N	10, 50	95,754	5.17	<.001
N: Linear trend	1, 5	353,923	23.07	<.001
Experiment 3b				
Training				
Block (B)	15, 345	326,798	32.77	<.001
Numerosity (N)	5, 115	1,329,747	101.22	<.001
B × N	75, 1,725	160,661	1.83	.050
N: Linear trend	1, 23	4,060,016	153.72	<.001
Transfer				
Stimulus type (ST)	2, 46	399,422	18.03	<.001
Numerosity (N)	5, 115	467,723	42.19	<.001
ST × N	10, 230	153,654	1.96	<.050
N: Linear trend	1, 23	1,634,946	58.07	<.001

Table A2 *Continued*

Effect	*df*	MS_e	*F*	*p*
		Experiment 4a		
Training				
Session (S)	3, 15	323,282	27.28	<.001
Numerosity (N)	5, 25	207,997	25.14	<.001
S × N	15, 75	53,960	5.90	<.050
N: Linear trend	1, 5	371,743	48.60	<.001
Transfer				
Stimulus type (ST)	2, 10	799,683	4.18	<.050
Numerosity (N)	5, 25	349,764	15.08	<.001
ST × N	10, 50	243,190	1.48	.174
N: Linear trend	1, 5	480,608	24.89	<.001
		Experiment 4b		
Training				
Block (B)	15, 315	469,673	13.51	<.001
Numerosity (N)	5, 105	737,400	157.60	<.001
B × N	75, 1,575	152,380	1.40	<.005
N: Linear trend	1, 21	1,870,818	284.50	<.001
Transfer				
Stimulus type (ST)	2, 42	228,249	14.69	<.001
Numerosity (N)	5, 105	254,707	87.23	<.001
ST × N	10, 210	174,157	2.11	<.050
N: Linear trend	1, 21	396,327	245.48	<.001

Received April 13, 1992
Revision received July 6, 1992
Accepted September 2, 1992 ■