Psychological Reports, 2006, 98, 419-426.

REPEATED QUESTIONING FOR ORDER OF EVENTS: DISPARATE EFFECTS FOLLOWING LOGICAL VERSUS RANDOM PRESENTATION[1]

MATTHEW R. KELLEY AND MELISSA S. LEHMAN

Lake Forest College

Summary.—In the present study, 72 college-age participants from an introductory psychology course viewed a series of 20 pictures depicting events surrounding a routine activity, i.e., eating at a cafeteria; these pictures were presented either in a logical order, e.g., enter cafeteria, pick up tray, stand in line, and select food, or in a random order. Three successive tests of free reconstruction of order indicated disparate effects of these conditions; random presentation produced significant forgetting of order information across tests, whereas logical presentation produced no change in performance across tests. Whereas randomly presented stimuli produced both reliable intertest recovery (reminiscence) and forgetting, neither result was observed following logical presentation. The implications of these data for eyewitness testimony for general theories of hypermnesia are discussed.

Ebbinghaus (1885/1964) was among the first to show that memory performance declines systematically with an increasing retention interval—performance declines rapidly at first, then gradually slows as performance reaches an asymptote. Since his time, a wealth of research has replicated this classic forgetting curve using a wide variety of designs and materials (for a review, see Rubin & Wenzel, 1996). Yet, forgetting is not the only outcome which can follow an increasing retention interval—under the appropriate conditions, memory performance can actually improve over time (e.g., Erdelyi & Becker, 1974). Often, when people are repeatedly questioned or tested, they may not remember information on an initial test but then will be able to remember the 'forgotten' information when queried on a later test. This phenomenon is known as reminiscence (e.g., Payne, 1987). Hypermnesia reflects a net improvement in memory performance across tests, and it occurs when the amount of intertest recovery (reminiscence) exceeds the amount of intertest forgetting (e.g., Payne, 1987).

Reminiscence appears to be a general characteristic of repeated testing, having been replicated with a variety of retention measures including free recall (e.g., Erdelyi & Becker, 1974; Roediger & Challis, 1989), cued recall (e.g., Otani & Hodge, 1991; Payne, Hembrooke, & Anastasi, 1993; Otani & Whiteman, 1994), recognition (e.g., Otani & Stimson, 1994), and free recon-

[1]Address correspondence to Dr. Matthew R. Kelley, Department of Psychology, Lake Forest College, Lake Forest, IL 60045 or e-mail (kelley@lfc.edu).

DOI 10.2466/PR0.98.2.419-426

struction of order (Kelley & Nairne, 2003). Hypermnesia, on the other hand, may depend, in part, on the specific retention measure employed in the study. Whereas free and cued recall typically yield hypermnesia (e.g., Payne, 1987; Otani & Whiteman, 1994), repeated recognition tests have produced either no change or an overall decline in performance across tests (e.g., Payne & Roediger, 1987; Otani & Hodge, 1991). Similarly, although Kelley and Nairne (2003) showed consistent reminiscence across successive tests of free reconstruction of order, overall retention for order declined.

The present study was designed to replicate and extend the findings of Kelley and Nairne (2003) by examining repeated testing for order information using more ecologically relevant stimuli, i.e., pictures of scenes, as opposed to using traditional laboratory-based materials, e.g., words. In the experiment, participants viewed a series of 20 pictures which depicted the events surrounding a routine activity, specifically that of selecting and eating lunch at a cafeteria. For half of the participants, these pictures were presented in a logical order, e.g., enter cafeteria, pick up tray, stand in line, and select food; the pictures were presented in a random order for the remaining participants, e.g., select food, pick up tray, stand in line, and enter cafeteria. Following the presentation of the final picture, participants received three consecutive tests of free reconstruction of order. In the free reconstruction task, participants were given the original pictures in a new random order and were asked to place the pictures back into their original order of appearance.

The random presentation condition was an analogue of Kelley and Nairne's first experiment (2003) wherein participants completed three successive free reconstruction tests following the presentation of random sets of either 25 words or 25 pictures of unrelated objects. Replication of these findings was expected, an overall drop in performance across tests despite reliable reminiscence with the present random condition. The logical presentation condition allowed examining how repeated testing might affect order retention for a real-life activity or event. One can generate a number of predictions regarding performance in this condition. First, previous research has shown a recall advantage for logically organized scripts of information as compared to the same information organized randomly (e.g., Bower & Clark-Meyers, 1980). Thus, it seems reasonable to assume that participants would perform better in the logical condition than in the random one.

With regard to how order retention would change across tests in the logical condition, one might assume that Kelley and Nairne's results (2003) reflect a general property of order retention and expect retention by order to decline across tests. However, the task demands involved in the logical condition seem quite different from those in the random condition. Retention of order in the logical condition might be guided by the use of script-related information. With such powerful retrieval cues in effect, one might expect high performance on an initial test with little or no change in performance across tests, i.e., no forgetting or reminiscence.

Method

Participants and Apparatus

Seventy-two college undergraduates of traditional college age participated for extra credit in an introductory psychology course. Groups of 12 or fewer participants were tested in sessions of approximately 45 min. Stimuli were presented and controlled with an IBM-compatible computer.

Materials and Design

Participants were assigned to one of two experimental conditions (logical or random presentation) or a control condition based on their order of arrival to the experiment. In all, there were 24 participants assigned to each of the three conditions. In both experimental conditions, participants viewed a set of 20 images. Each image depicted a specific scene that could occur while in a cafeteria, e.g., enter cafeteria, pick up tray, stand in line. In the Logical condition, the 20 images were displayed in an order consistent with a typical visit to the cafeteria. In the Random condition, the 20 images were presented to the participants in a random order. Within each condition, images were presented in exactly the same order to all participants. Following presentation, participants received the same test sheet for reconstruction of order, which consisted of a column of 20 blanks located on the left margin, and numbered 1 through 20, and thumbnail photographs of the 20 images labeled A through T, randomly organized into four columns of five pictures each, positioned to the right of the numbers.

In the Control condition, participants were asked to reconstruct the pictures in a logical sequence without having participated in the presentation phase; these participants completed only one reconstruction test. Although there are many plausible, logical sequences of the 20 images, the Control condition was included to assess the likelihood that participants would generate, by chance, the one sequence employed in the Logical condition. Without this Control condition, one could not specify whether participants in the Logical condition actually completed the task by memory of the original sequence or whether they could simply order the pictures based on a general script for a cafeteria visit.

Procedure

Participants in the experimental conditions were asked to remember the set of 20 images for a subsequent memory-reconstruction task. They were instructed about the exact nature of the memory test prior to list presentation, but they only expected one test. The stimuli were projected onto a screen at a rate of one item every 5 sec. Participants were instructed to pay close attention to each image as it was presented. Following the final item, participants received the first test sheet and were asked to place the images into their original order of appearance by writing the letter of each image (A-T) into its appropriate ordinal position (1–20) within 7 min. Control participants were simply given the reconstruction test sheet and were instructed to reorder the images so that they represented a typical visit to a cafeteria.

Following the first test, the participants' test sheets were collected and the Control participants were then dismissed from the experiment. Participants in the experimental conditions, however, were asked to complete a second test, identical to the first. They were encouraged to try to improve their performance on the second test relative to the first one. The second 7-min. test began immediately following the instructions. Upon completion of Test 2, participants were given more instructions and the third and final test; the general procedure of Test 3 was similar to that of Test 2. Again, participants were encouraged to improve their performance on this test relative to the first two.

The reconstruction of order tests were examined using two different scoring criteria, strict order and relative order. According to the strict-order method, an item was marked as correct only if it was reconstructed in its original serial position. Although the strict-order method is the standard scoring criterion employed in reconstruction tests, the technique is rather unforgiving in its estimate of order retention. Imagine that a participant misplaces one item in a sequence, e.g., 1-3-4-5-6-7-8-9-10-2. In this situation, the participant's strict order accuracy is only 10%, which masks the fact that Items 3 to 10 were reconstructed in their appropriate serial order.

To gather a more complete depiction of retained order, the data were also examined using a relative-order scoring method. With this technique, an item was scored as correct if the items immediately preceding and immediately following were presented relatively before and after. By using the relative-order method, the previous participant's accuracy improved to 80%, which better reflected the fact that the middle items were properly ordered. Specifically, all of the reconstructed items are counted as correct except for Items 10 and 2. Item 10 is scored as incorrect because the item that immediately follows 10 is Item 2, which originally appeared earlier in the list than did Item 10. Similarly, Item 2 is incorrect because the immediately preceding item is 10, which originally appeared later than Item 2 in the list.

Results

Table 1 displays the strict- and relative-order net reconstruction scores for each test in each experimental and control condition. Separate 2 (presentation order: logical, random) × 3 (test number: 1, 2, 3) mixed-factor analyses of variance were performed on the mean number of correctly reconstructed items for the strict-order and the relative-order data, respectively. As expected, both analyses indicated significant main effects of presentation order ($F_{1,46} = 277.91$, $p < .001$; $F_{1,46} = 269.70$, $p < .001$), where the Logical conditions yielded greater reconstruction accuracy than did the Random conditions. Reliable main effects of test number ($F_{2,92} = 12.08$, $p < .001$; $F_{2,92} = 14.40$, $p < .001$), were also noted, suggesting that overall reconstruction accuracy declined as subsequent tests were given. In addition, the analysis of variance indicated significant interactions between presentation order and test number ($F_{2,92} = 4.78$, $p < .05$; $F_{2,92} = 3.20$, $p < .05$).

TABLE 1
STRICT- AND RELATIVE-ORDER NET RECONSTRUCTION SCORES FOLLOWING EACH TEST FOR PICTURES IN EACH EXPERIMENTAL AND CONTROL CONDITION

Dependent Measure	Logical	Random	Control
Strict-order Scoring			
Test 1	.93	.29	.34
Test 2	.92	.23	
Test 3	.90	.20	
Change ($T_3 - T_1$)	–.03	–.09	
Relative-order Scoring			
Test 1	.95	.44	.59
Test 2	.95	.39	
Test 3	.93	.35	
Change	–.02	–.09	

A series of Newman-Keuls tests were employed to examine these interactions further. With both strict- and relative-order scoring, the analyses suggested that the findings from the Random condition were consistent with those reported by Kelley and Nairne (2003). Net reconstruction performance declined significantly across tests ($T_1 > T_2 > T_3$; $ps < .05$). Of the 24 participants in the Random reconstruction condition with strict-order scoring, 17 showed a drop in performance from the first to the last test, 2 showed hypermnesia, and there were 5 ties. The pattern was similar following relative-order scoring (19 declined, 4 increased, and 1 tie). Although the data from the Logical condition indicated a small numerical drop in performance across tests, this decrease was not statistically significant with either scoring criterion ($T_1 = T_2 = T3$; $ps > .05$). Of the 24 participants in the logical reconstruction condition with strict-order scoring, 18 showed no change in performance from the T_1 to T_3, 6 showed forgetting, and none showed hypermnesia; the relative-order scoring showed exactly the same pattern of scores. Thus, the interactions indicated disparate effects of the Logical and Random conditions. Random presentation produced reliable forgetting across tests, whereas Logical presentation produced no change in performance across tests.

With regard to the intertest forgetting and recovery data (see Table 2), findings in the Random condition supported those described by Kelley and Nairne (2003); both studies indicated consistent intertest recovery rates of about 5% across both early and late tests, and intertest forgetting which exceeded intertest recovery. Given the pervasiveness of reminiscence in the repeated testing literature, it is interesting that participants in the Logical condition did not produce reliable recovery across the tests. Performance in this condition, however, was near the ceiling and thus did not offer much opportunity for recovery. One potential explanation for the superior performance in the Logical condition is that powerful script-related cues helped to guide retrieval in this task. Accordingly, one might expect little (if any) retrieval variability and, hence, little difference in reminiscence and forgetting between tests.

TABLE 2
RATES OF INTEREST RECOVERY AND FORGETTING BETWEEN EARLY AND LATE TESTS FOR PICTURES IN BOTH EXPERIMENTAL CONDITIONS (STRICT-ORDER SCORING ONLY)

Dependent Measure	Logical	Condition
Recovery		
Early ($T_1:T_2$)	.01	.05
Late ($T_2:T_3$)	.00	.05
Forgetting		
Early ($T_1:T_2$)	.01	.11
Late ($T_2:T_3$)	.02	.08

If one assumes that script-related cues contribute to the elevated performance in the logical condition, then one could question whether mnemonic processes are involved in this task. In other words, one could argue that, during the test phase, participants might simply order the pictures based on a general script for a cafeteria visit and may not even consult their memories for the actual events. In this situation, one would expect the control participants, who were asked to reconstruct the pictures into a logical sequence without having participated in the presentation phase, and the logical participants to produce equivalent performance. Two separate independent-sample *t* tests examined this possibility by comparing reconstruction performance on the first test of the Logical and Control groups. The *t* ratios indicated that Logical performance (.93; .95) was significantly higher than Control performance (.34; .59) for strict-order and relative-order scoring criteria ($t_{46} = 9.11$, $p < .001$; $t_{46} = 7.86$, $p < .001$). The 24 Control participants generated 23 unique sequences of events, and none matched the original sequence employed in the Logical condition. Clearly, by relying solely on their script-based knowledge for cafeteria visits, participants accurately reconstructed portions of the relative order of the original sequence. To achieve the near ceiling performance displayed in the Logical condition, however, participants presumably used a combination of these script-based cues along with their memories for the actual event.

General Discussion

The present experiment established that the effects of repeated testing on retention for order depend critically on the presentation order of the information to be remembered. Following a set of randomly ordered pictorial stimuli, participants' net retention for order was poor and then declined significanatly across successive free reconstruction of order tests. In contrast, when the same stimuli were presented in a logical order, net reconstruction performance was near the ceiling and did not change over tests. Also, whereas the randomly presented stimuli produced both reliable intertest recovery (reminiscence) and forgetting, neither result was seen in the logical presentation condition.

It is plausible that this benefit of logical ordering might stem from the use of script- or schema-based retrieval processes (e.g., Schank & Abelson, 1977). Scripts are knowledge representations that consist of information about commonly experienced events, such as information about the order in which events generally occur (Schank & Abelson, 1977). It seems likely that, when a script becomes activated in memory, participants' retrieval may be guided by powerful script cues. In these situations, one might expect relatively stable retention for order across repeated tests with very little variability of retrieval. However, to the extent that events do not activate or are inconsistent with a pre-existing schema, one might expect an increase in retrieval variability and perhaps a decline in retention of order over tests, as seen following the presentation of randomly ordered information.

These ideas of retrieval variability appear to be consistent with aspects of the retrieval dynamics explanation of reminiscence and hypermnesia (Payne, *et al.*, 1993). Payne, *et al.* (1993) argued that retrieval in tasks such as free and cued recall is driven by fluctuating cues, e.g., context cues, item cues, and category cues. Greater fluctuation in cues should foster greater variability of retrieval, which in turn, should elevate intertest recovery and forgetting in these tasks. Retrieval cues in other tasks such as recognition are not as likely to fluctuate because the cues provided in a recognition test are the target items. Thus, there should be little, if any, recovery or forgetting across tests. One can extend their reasoning to account for the findings in the Logical condition by assuming that script-based retrieval cues are unlikely to vary over tests. An absence of fluctuating cues should lead to absences of intertest recovery and forgetting.

Although one can apply some of the retrieval dynamics ideas to the current study, in its present form, a retrieval dynamics approach does not explicitly address retention of order information and therefore does not account for the full range of data presented here. This lack also is present in the other leading theories of reminiscence and hypermnesia (i.e., the cumulative recall hypothesis: Roediger, 1982; Roediger, Payne, Gillespie, & Lean 1982; the relational/item-specific account: McDaniel, Moore, & Whiteman, 1998), which were designed primarily to explain the recovery and loss of item, as opposed to order, information. To provide complete explanations of reminiscence and hypermnesia, these theories need to account for the disparate effects of repeated testing following logical and random presentation as well as the data reported by Kelley and Nairne (2003); the required modifications are not readily apparent and these explanations are likely to require substantial revision.

REFERENCES

Bower, G. H., & Clark-Meyers, G. (1980) Memory for scripts with organized vs randomized presentations. *British Journal of Psychology*, 71, 369-377.

Ebbinghaus, H. (1885/1964) *Memory: a contribution to experimental psychology.* (H. A. Ruger & C. E. Bussenius, Transl.) New York: Dover.

Erdelyi, M. H., & Becker, J. (1974) Hypermnesia for pictures: incremental memory for pictures but not for words in multiple recall trials. *Cognitive Psychology*, 6, 159-171.

Kelley, M. R., & Nairne, J. S. (2003) Remembering the forgotten? Reminiscence, hypermnesia, and memory for order. *Quarterly Journal of Experimental Psychology*, 56A, 577-599.

McDaniel, M. A., Moore, B. A., & Whiteman, H. L. (1998) Dynamic changes in hypermnesia across early and late tests: a relational/item-specific account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 173-185.

Otani, H., & Hodge, M. H. (1991) Does hypermnesia occur in recognition and cued recall? *American Journal of Psychology*, 104, 101-116.

Otani, H., & Stimson, M. J. (1994) A further attempt to demonstrate hypermnesia in recognition. *The Psychological Record*, 44, 25-34.

Otani, H., & Whiteman, H. L. (1994) Cued recall hypermnesia is not an artifact of response bias. *American Journal of Psychology*, 107, 401-421.

Payne, D. G. (1987) Hypermnesia and reminiscence in recall: a historical and empirical review. *Psychological Bulletin*, 101, 5-27.

Payne, D. G., Hembrooke, H. A., & Anastasi, J. S. (1993) Hypermnesia in free and cued recall. *Memory & Cognition*, 21, 48-62.

Payne, D. G., & Roediger, H. L. III. (1987) Hypermnesia occurs in recall but not in recognition. *American Journal of Psychology*, 100, 145-165.

Roediger, H. L. III. (1982) Hypermnesia: the importance of recall time and asymptotic recall level. *Journal of Verbal Learning and Verbal Behavior*, 21, 662-665.

Roediger, H. L. III, & Challis, B. H. (1989) Hypermnesia: improvements in recall with repeated testing. In C. Izawa (Ed.), *Current issues in cognitive processes: the Tulane Flowerree Symposium on Cognition*. Hillsdale, NJ: Erlbaum. Pp. 175-199.

Roediger, H. L. III, Payne, D. G., Gillespie, G. L., & Lean, D. S. (1982) Hypermnesia as determined by level of recall. *Journal of Verbal Learning and Verbal Behavior*, 21, 635-655.

Rubin, D. C., & Wenzel, A. C. (1996) One hundred years of forgetting: a quantitative description of retention. *Psychological Review*, 103, 734-760.

Schank, R. C., & Abelson, R. P. (1977) *Scripts, plans, goals and understanding: an inquiry into human knowledge structures*. Oxford, Eng.: Erlbaum.

Accepted March 2, 2006.