

Testing protects against proactive interference in face–name learning

Yana Weinstein · Kathleen B. McDermott · Karl K. Szpunar

Published online: 5 April 2011
© Psychonomic Society, Inc. 2011

Abstract Learning face–name pairings at a social function becomes increasingly more difficult the more individuals one meets. This phenomenon is attributable to proactive interference—the negative influence of prior learning on subsequent learning. Recent evidence suggests that taking a memory test can alleviate proactive interference in verbal list learning paradigms. We apply this technique to face–name pair learning. Participants studied four lists of 12 face–name pairings and either attempted to name the 12 faces just studied after every list or did not. Recall attempts after every list improved learning of the fourth list by over 100%. Moreover, no reduction in learning of face–name pairings occurred from list 1 to list 4 for participants who attempted to name studied faces between lists. These results suggest that testing oneself on the names of a group of new acquaintances before moving on to the next group is an effective mnemonic technique for social functions.

Keywords Face–name learning · Proactive interference · Testing

Learning the names of new acquaintances in a novel context represents a real-world memory task often

Support for this research was provided by a James S. McDonnell Foundation 21st Century Science Initiative grant: Bridging Brain, Mind and Behavior/Collaborative Award.

Y. Weinstein (✉) · K. B. McDermott
Department of Psychology, Washington University in St. Louis,
Campus Box 1125, One Brookings Drive,
St. Louis, MO 63130, USA
e-mail: y.weinstein@wustl.edu

K. K. Szpunar
Department of Psychology, Harvard University,
Cambridge, MA, USA

approached with a distinct sense of pessimism. At first, learning the names of a few individuals may not seem too daunting. However, as one is introduced to more groups of individuals, the prospect of successfully learning new names seems futile. In fact, at some point remembering any additional names becomes unlikely because there are already so many other face–name pairings of which to keep track. This often leaves one vulnerable to social tensions that arise if a particular person’s name is not remembered in the future (or worse yet, if the wrong name is generated).

A considerable amount of work has focused on learning novel face–name pairings (e.g., Carpenter & DeLosh, 2005; Helder & Shaughnessy, 2008; Landauer & Bjork, 1978; Morris, Fritz, Jackson, Nichol, & Roberts, 2005), and some mnemonic techniques have been suggested that enhance memory of specific face–name pairs (Carney, Levin, & Stackhouse, 1997; Morris et al., 2005; Neuschatz, Preston, Togliola, & Neuschatz, 2005). However, there has been surprisingly little consideration about why learning names becomes more difficult over an extended period of time, such as occurs when one attends a party. In this article, we present a brief overview of the cognitive phenomenon thought to impair learning over time, explain how this phenomenon applies to learning face–name pairings, and present a strategy for greatly enhancing the efficiency with which face–name pairings are learned over time that can easily be applied in real-world settings.

Learning and proactive interference

Generally, the longer one is engaged in the study of a particular set of to-be-learned materials (e.g., lists of words, sets of class notes, etc.), the more difficult it becomes to learn those materials over time. This striking limitation of

learning is the result of a cognitive phenomenon known as proactive interference (PI)—the interference of previously learned materials with the learning of subsequent materials. For instance, Underwood (1957) demonstrated that, when learning multiple lists of words in a prolonged study sequence, participants do not learn lists of words encountered later on in the sequence as well as lists of words encountered earlier on (see also Postman & Keppel, 1977).

Although this learning phenomenon has been studied using simple laboratory materials (e.g., lists of words), we hypothesize that PI also places important limitations on the learning of face–name pairings. Indeed, the act of learning face–name pairings can be thought of as a complex form of list learning. Nonetheless, there exists little experimental evidence that addresses the relation between PI and face–name learning. A few studies have addressed the role of interference in face–name learning, but these studies have focused on retroactive interference (i.e., the effect of learning new information on the retrieval of previously learned information; Deffenbacher, Carr, & Leu, 1981). Bengner et al. (2006) did investigate PI in face–name learning in epileptics, but this study used face–name pair recognition instead of recall, which does not accurately mirror a real-life situation where you would be expected to retrieve, rather than verify, a name.

Alleviating PI

It has recently been demonstrated that PI can be alleviated via the simple act of testing one's memory. Szpunar, McDermott, and Roediger (2008) asked students to study five lists of words. After studying each list (and completing a 60-s math distractor), students completed one of two additional tasks before moving onto the next list: (1) a further 60 s of math or (2) a free recall test for the recently studied list (also lasting 60 s). Students were told that the program would randomly determine the occurrence of additional math/recall after each list and that they should study all lists expecting a test. Students were also told that there would be a final memory test on all of the words from all of the lists at the end of the experiment (regardless of how many lists were tested). In reality, half of the students were tested on the first four lists (tested group), and half were not tested on the first four lists (untested group). However, all students were tested on the fifth list and received a cumulative test.

Relative to students in the untested group, students in the tested group were twice as likely to recall words from the fifth list. Moreover, students in the tested group almost never generated prior list intrusions (i.e., words from lists 1–4) when recalling the fifth list, whereas students in the untested group generated just as many prior list intrusions as list 5 words

(even though they had explicitly been instructed to recall words only from the list they had just studied). This benefit of prior testing on learning of list 5 words was also apparent in the final recall test (see also Darley & Murdock, 1971; Pastötter, Schicker, Niedernhuber, & Bäuml, 2011; Tulving & Watkins, 1974).

PI and cued recall

Although face–name learning has often been investigated in its own right, this type of learning can essentially be thought of as a paired associates task with more complex materials. A large body of literature exists on PI with cued-recall (e.g., Underwood, 1957). However, this field has utilized a paradigm that differs from face–name pair learning in one important way: The standard cued recall paradigm involves repeated cues (i.e., the much-used A–B, A–C paradigm), and research into PI using this paradigm has focused on whether learning one cue–target associations impairs learning of a subsequent association of the same cue with a different target. In face–name learning, on the other hand, every cue is a unique face.¹ According to the classic response competition explanation, unique cues should eliminate PI (McGeoch, 1942). For this reason, the A–B, C–D (i.e., unique cue–target association) version of the cued recall task is usually included in interference studies as the control condition, where little or no PI occurs.

Could it be that cued recall performance with unique cue–target pairs does not suffer from the buildup of PI across multiple lists? In that case, taking a test between lists to eliminate PI would not be necessary. Delving further into the classic interference literature, however, we find that the specific paradigms used to study PI almost always involved immediate tests after each studied list (usually two) to get performance up to a criterion, and only then was interference measured on further tests (e.g., Postman & Keppel, 1977). One demonstration that PI does build up with unique cue–target pairs comes from an unpublished dissertation by Malis (1970), later described in Tulving and Watkins (1974). In the part of his experiment relevant to the present discussion, Malis had participants study either one list of cue–target pairs or two lists separated by a 30-s interval, where none of the cues or targets in list 2 overlapped with those in list 1. Both groups then received a cued recall test on list 2; participants who had previously learned list 1 performed worse (mean accuracy of .50) than participants who learned only list 2 (mean accuracy of .68). The finding that PI builds up even with unique cue–target associations

¹ This issue is somewhat complicated by the fact that in real life, the *targets* (names) do sometimes repeat. However, for the purpose of this study, we are interested in the learning of unique face–name pairings.

is inconsistent with the response competition explanation of PI but is consistent with a more general view that attributes PI to the inability to differentiate between materials learned at different points in time (e.g., Crowder, 1976).

Present experiment

The experiment presented in this article first tested the assumption that PI hampers face–name learning even though each face provides a unique cue at retrieval. If this assumption held true, our next aim was to determine whether this buildup of PI could be alleviated by interpolated tests, as it can in the free recall paradigm (Szpunar et al., 2008). If presenting intermittent tests between sets of materials has a beneficial effect on learning face–name pairings over time, such a technique could easily be applied to the improvement of face–name learning in everyday situations.

In the present study, we adapted the Szpunar et al. (2008) PI paradigm to the learning of face–name pairings. In doing so, we made two important changes to the paradigm. First, participants studied lists of face–name pairings rather than words. Second, the tests that participants received were cued recall, rather than free recall. For each test, participants were presented with a series of faces and were asked to generate the associated names. These procedural changes allowed us to model an everyday learning situation while addressing the theoretical question outlined above with regard to the process involved in the reduction of PI through interpolated testing.

Method

Participants Thirty-two Washington University undergraduates participated in the experiment for course credit or payment. Half the participants were randomly assigned to the untested group, and the other half to the tested group.

Materials Four study lists of 12 face–name pairs were constructed. Faces were pictures of males from the Psychological Image Collection at Stirling database (<http://pics.psych.stir.ac.uk/>). One of 48 popular male first names was randomly assigned to each face; name assignments were consistent between participants.

Design Testing group (tested/untested) was manipulated between participants. The order of the four study lists was counterbalanced using a Latin square design: 4 participants in each testing group studied the lists in one of the four possible list orders. Correct recall, prior list intrusions, and current list intrusions were analyzed for each initial test

taken. That is, for the tested group, performance on lists 1–4 was examined, whereas for the untested group, only performance after list 4 was analyzed. In addition, correct recall and intrusions on the final cumulative test were compared between the two testing groups.

Procedure Participants were told that they would study four lists of face–name pairs and that, after the presentation of each individual list, they might or might not perform math problems and might or might not be asked to name the faces from the immediately preceding list. They were informed that the occurrence of a test after each list was determined randomly by the program. In reality, there were only two testing schedules: either 60 s of math followed by a test after every list (tested group) or 160 s of math after lists 1–3 and 60 s of math plus a test after the fourth list (untested group). Participants were also warned that they would be asked to name the faces from all the studied lists on a final cumulative test.

At study, face–name pairs were presented in random order for 4,000 ms each, with a 500-ms inter-stimulus interval. Names appeared in large font below faces. After each of the first three lists, all participants completed math problems for 60 s. Subsequently, participants either took a cued recall test that lasted 100 s (tested group) or completed math problems for the same amount of time (untested group). After the fourth list, all participants took the cued recall test following 60 s of math problems. In these initial cued recall tests, participants were presented with the faces from the previous list in random order for 8,000 ms each and were asked to type in the name paired with the face at study. After the fourth list test, participants completed a cumulative cued recall test that probed their memory for all the faces they had encountered. On this test, all the studied faces were presented in random order, and participants attempted to recall the name that had been paired with each face at study, in their own time, and could choose not to respond to some faces.

Results

Scoring Responses were scored as correct if they were the exact assigned name of the face or a close misspelling. For instance, “Arron” was accepted in place of “Aaron.” There was a difference in scoring of the initial list tests and the cumulative test because the former were time constrained, whereas the latter was not. For the initial list tests, if participants omitted the last few letters of a name because they were in the middle of typing their response when the 8,000-ms deadline terminated the trial, their response was accepted as correct. For the cumulative test, this additional

scoring rule was not necessary, since participants had unlimited time on each trial.

Initial list tests The top row of Table 1 presents the number of correctly named faces for the tests given after lists 1–4 in the tested group. The table shows that participants in the tested group consistently named over half of the faces on each test correctly ($M = 6.78$ across the four lists). The first two bars of Fig. 1 compare the number of correctly named faces on the fourth test between the tested and untested groups. Participants in the untested group, contrary to the tested group, correctly named fewer than a third of the faces on the fourth list test. Analogously, the second row of Table 1 presents the number of prior list intrusions in the tested group for each of lists 2–4 and shows that participants in the tested group produced almost no such intrusions (an average of $M = 0.19$ prior list intrusions per list). Looking at Fig. 1, on the other hand, we see that participants who had not been tested on every list produced over 10 times as many intrusions on the fourth list test ($M = 2.44$ intrusions). These comparisons were supported by statistical analyses: A within-subjects ANOVA showed no effect of list number on correct recall for lists 1–4 in the tested group, $F < 1$, $p = .64$, whereas an independent-samples t -test showed a significant difference between the tested and untested groups in correct recall on the fourth list test, $t(30) = 5.03$, $d = 1.78$. Similarly, there was no difference in prior list intrusions across lists 2–4 for the tested group, $F < 1$, $p = .70$, but a significant difference between the tested and untested groups in the number of prior list intrusions produced on the list 4 test, $t(30) = 4.64$, $d = 1.64$.

The pattern was somewhat different for current list intrusions, shown in the third row of Table 1 as a function of list for the tested group; current list intrusions on list 4 for the tested and untested groups are compared in Fig. 1. As can be seen in Table 1, participants in the tested group made more current list intrusions on earlier lists and

Table 1 Mean correctly named faces, prior list intrusions, current list intrusions, extra-list intrusions, and omissions on tests taken after lists 1–4 for the tested condition (*SE* in brackets)

Response Category	List 1	List 2	List 3	List 4
Correct	6.19 (0.71)	6.88 (0.83)	7.00 (0.52)	7.06 (0.62)
Prior list intrusions	–	0.13 (0.09)	0.25 (0.11)	0.19 (0.14)
Current list intrusions	2.06 (0.37)	1.19 (0.31)	0.50 (0.16)	0.56 (0.20)
Extra-list intrusions	0.88 (0.35)	0.75 (0.39)	0.44 (0.26)	0.31 (0.15)
Omissions	2.88 (0.45)	3.06 (0.57)	3.81 (0.84)	3.88 (0.50)
Total	12	12	12	12

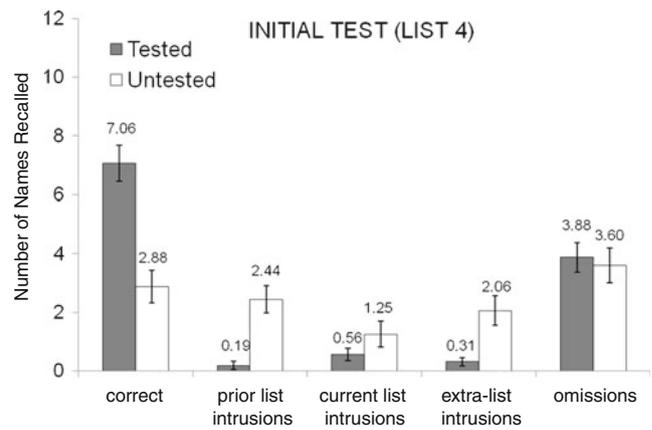


Fig. 1 Number of correctly named faces, prior list intrusions, current list intrusions, extra-list intrusions, and omissions on the fourth list test as a function of testing conditions (tested = previously tested on lists 1–3; untested = tested only on list 4). Values for each of the five categories sum to 12 for each condition. Error bars represent *SEM*

improved gradually across the four lists. In addition, participants in the untested group made roughly the same number of current list intrusions on the fourth list test as the tested group (see Fig. 1, third pair of bars from the left). A significant effect of list number on current list intrusions confirmed the observed data pattern for the tested group, $F(3, 45) = 9.74$, $\eta_p^2 = .39$; the comparison between tested and untested groups for current list intrusions on list 4 was not significant, $t(30) = 1.41$, $p = .17$.

The remaining responses were made up of extra-list intrusions (names that had not been studied on any list) and omissions. Although extra-list intrusions numerically declined across lists 1–4 in the tested group (see the penultimate row of Table 1), this trend did not approach significance ($p = .23$). However, as can be seen in the penultimate pair of bars in Fig. 1, participants in the untested group produced more than 6 times the number of extra-list intrusions on list 4, as compared with the tested group, $t(30) = 3.38$, $d = 1.19$. Omissions, on the other hand, numerically increased across lists 1–4 for the tested group (see the bottom row of Table 1), but once again this trend was not significant ($p = .18$). In addition, participants in the tested and untested groups did not differ on this measure on the list 4 test (rightmost pair of bars in Fig. 1; $p = .53$). That is, participants in the two groups attempted to name the same proportion of faces (approximately 70% of the faces).

Cumulative test The number of faces named correctly on the cumulative test is presented in Fig. 2a. Participants in the tested group correctly named almost quadruple the number of faces ($M = 21.88$ faces) that participants who took a test only after the fourth list were able to name correctly ($M = 5.94$ faces), $t(30) = 7.84$, $d = 2.78$. Figure 2a also shows the number of incorrectly named faces; this

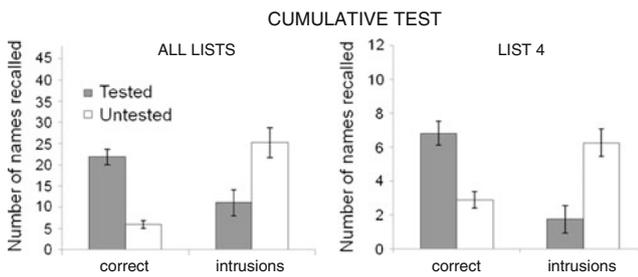


Fig. 2 Left panel: Number of correctly named faces and intrusions on the cumulative test for the tested and untested groups, out of a possible 48. Right panel: Number of correctly named faces and intrusions on the cumulative test for faces that had appeared in the fourth list, out of a possible 12

includes any face that was given the name of another face from the same list, the name of a face from a different list, or a name that had not been presented at study. These data demonstrate that participants in the untested group attempted to name as many faces as those in the tested group but, for the most part, did so inaccurately. Participants in the untested group incorrectly named more than double the number of faces, as compared with participants in the tested group, $t(30) = 3.04$, $d = 1.08$. Figure 2b further confirms that the same pattern was found if only faces from the fourth list were analyzed: Even though both groups of participants received an initial test on this list, those in the tested group correctly named more faces, $t(30) = 4.60$, $d = 1.63$, and those in the untested group incorrectly named more faces, $t(30) = 3.97$, $d = 1.40$.

Discussion

Using an adapted version of Szpunar et al.'s (2008) PI paradigm, we showed that performance on a cued recall task with unique cue–target (face–name) associations is negatively impacted by the buildup of PI. Specifically, participants who were tested only on the final study list were less likely than participants who had been tested on every list to name faces correctly and were more likely to use names from previous lists. Further replicating and extending the results in Szpunar et al., we also found that participants who had been tested on each study list performed equally well on each of the four lists (and produced equally few prior list intrusions on lists 2–4). In short, testing was found to protect participants against the negative influence of PI that typically builds up over the course of a study session.

Two candidate theories have been proposed to explain why intermittent testing during acquisition protects the learner from PI. Szpunar et al. (2008) claimed that taking a test after each list allows for better list discrimination. According to this explanation, participants who have been

tested on each list are better able to limit their search (during retrieval) to the most recent list by segregating previously tested information from as-of-yet untested information (cf. Johnson, Hashtroudi, & Lindsay, 1993). A different conceptualization, proposed by Pastötter et al. (2011), is that intermittent testing helps to maintain the efficiency with which lists of words are encoded in the course of a study sequence. Using EEG recordings during encoding, Pastötter et al. demonstrated an increase in alpha power (thought to indicate failure of attention) across encoding of multiple lists with no interpolated testing and no such alpha power increase with testing. On the basis of this finding, the authors hypothesized that testing “resets” encoding to optimum efficiency for the next list.

While our experiment was not designed to distinguish between these possible explanations, our cued recall procedure allowed us to more closely examine the nature of the errors made in the tested and untested groups on the fourth list test. First, we found that the proportion of omissions was equal in the two groups, which conceptually replicates Szpunar et al.'s (2008) finding that participants in the two conditions produced the same total number of words in the free recall version of the task (with the relative proportion correct vs. prior list intrusions driving differences between conditions). Second, in a follow-up analysis, we looked at the proportion of commission errors on the list 4 test that came from the current list for each of the two groups. In the tested group, .53 of commission errors were current list intrusions, whereas for the untested group, this proportion was only .22. This comparison suggests that participants in the tested condition were better able to constrain recall to the most recent list than were participants in the untested group, consistent with the proposal by Szpunar and colleagues. On the other hand, the lower proportion of current list intrusions in the untested group could also be explained by poorer encoding of that list, consistent with the proposal by Pastötter et al., (2011). In short, the data here are consistent with both frameworks, which are not mutually exclusive. That is, intermittent testing may “reset” encoding efficiency and may also help delimit the search set during retrieval.

In addition to reducing PI, testing produced the standard benefit noted in recent work on the testing effect (e.g., Roediger & Karpicke, 2006): The tested group performed 4 times better on the cumulative test than did the untested group. Our results add to the body of evidence supporting the role of testing in promoting long-term retention of face names (e.g., Carpenter & DeLosh, 2005; Morris et al., 2005). However, the superior cumulative test performance in the tested group cannot be attributed solely to the testing effect: Superior performance was also observed on the faces studied in the fourth list—those that had been initially tested for both groups. Testing is a far easier technique to implement than some of the more complicated mnemonic strategies favored

in the memory improvement literature (e.g., McCarty, 1980), and our study shows that testing can ensure that one has an equal chance of learning the names of members from multiple groups at a social function. Morris et al. already demonstrated that covert retrieval of new acquaintances' names at a party can improve long-term retention of names. That is, it is not necessary to say newly learned names aloud in order to benefit from the testing effect, making this a discreet strategy to employ in public. Combining this evidence for the effectiveness of covert self-testing with our new finding that testing protects against PI in face-name learning, self-testing on the names of one group before moving on to the next would be a very effective strategy to engage in at social functions.

References

- Bengner, T., Malina, T., Lindenau, M., Voges, B., Goebell, E., & Stodieck, S. (2006). Epilepsy increases vulnerability of long-term face recognition to proactive interference. *Epilepsy & Behavior*, *8*, 220–227.
- Carney, R. N., Levin, J. R., & Stackhouse, T. L. (1997). The face-name mnemonic strategy for a different perspective. *Contemporary Educational Psychology*, *22*, 399–412.
- Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, *19*, 619–636.
- Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale: Erlbaum.
- Darley, C. F., & Murdock, B. B., Jr. (1971). Effects of prior free recall testing on final recall and recognition. *Journal of Experimental Psychology*, *91*, 66–73.
- Deffenbacher, K. A., Carr, T. H., & Leu, J. R. (1981). Memory for words, pictures, and faces: Retroactive interference, forgetting, and reminiscence. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 299–305.
- Helder, E., & Shaughnessy, J. J. (2008). Retrieval opportunities while multitasking improve name recall. *Memory*, *16*, 896–909.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.
- Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruenberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). New York: Academic Press.
- Malis, G. (1970). *The priority effect in the A-B, A-C paradigm*. Unpublished doctoral dissertation, University of Toronto.
- McCarty, D. L. (1980). Investigation of a visual imagery mnemonic device for acquiring face-name associations. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 145–155.
- McGeoch, J. A. (1942). *The psychology of human learning: An introduction*. New York: Longmans.
- Morris, P. E., Fritz, C. O., Jackson, L., Nichol, E., & Roberts, E. (2005). Strategies for learning proper names: Expanding retrieval practice, meaning, and imagery. *Applied Cognitive Psychology*, *19*, 779–798.
- Neuschatz, J. S., Preston, E. L., Togli, M. P., & Neuschatz, J. S. (2005). Comparison of the efficacy of two name-learning techniques: Expanding rehearsal and name-face imagery. *The American Journal of Psychology*, *118*, 79–102.
- Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *37*, 287–297.
- Postman, L., & Keppel, G. (1977). Conditions of cumulative proactive inhibition. *Journal of Experimental Psychology: General*, *106*, 376–403.
- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255.
- Szpunar, K. K., McDermott, K. B., & Roediger, H. L., III. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1392–1399.
- Tulving, E., & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior*, *13*, 181–193.
- Underwood, B. J. (1957). Interference and forgetting. *Psychological Review*, *64*, 49–60.