Do distractors interfere with memory for study pairs in associative recognition?

PIERRE PERRUCHET, ARNAUD REY, and EIMERIC HIVERT
University of Bourgogne, Dijon, France

and

SÉBASTIEN PACTON
University René Descartes, Paris, France

In an associative recognition task, distractors generally consist of a rearrangement of the items composing the study pairs. This makes it possible that processing the distractors generates retroactive interference on memory for the study pairs. In Experiment 1, we explored this possibility in a *yes/no* recognition test concerning previously learned arbitrary associations between visual symbols and auditory syllables. Rearranged pairs had a deleterious impact on the accuracy and the speed of responses to related correct pairs. This effect did not vary as a function of the number of training blocks, and furthermore, in Experiment 2, the same effect was observed for overlearned small multiplication facts. These results suggest that exposure to potentially confounding information generates interference even if this information is known to be incorrect. Some implications are outlined, especially with regard to the widespread use of multiple-choice tests in knowledge evaluation.

In research on associative recognition (e.g., Cleary, Curran, & Greene, 2001; Hockley, 1991; Light, Patterson, Chung, & Healy, 2004; Westerman, 2001), participants first study pairs of items (e.g., *AB*, *CD*) and then have to discriminate between intact (e.g., *AB*) and rearranged (e.g., *AD*) pairs. It is worthy of note that the rearranged pairs, or *distractors*, are patterned in a way similar to that of the interfering information in classical research on retroactive interference using paired-associate paradigms. In that context, participants are first trained with *AB*, then with *AD*. A vast amount of research has shown that processing *AD* has a deleterious effect on the subsequent retrieval of *AB* (see Anderson & Neely, 1996, for a review). This parallel between associative recognition and retroactive interference paradigms raises the possibility that processing the rearranged pairs in associative recognition tasks could be detrimental to the subsequent correct endorsement of the intact pairs. However, this possibility requires empirical confirmation, because participants in associative recognition studies are not asked to learn the *AD* pairs and, furthermore, they presumably appraise some of the *AD*

pairs as incorrect. Also, the standard paired-associate paradigms require the retrieval of a response (e.g., *B*) given a stimulus (e.g., *A*), whereas a judgment on displayed pairs of events (e.g., *AB*) is asked for in associative recognition paradigms. Likewise, the ordering and the timing of the *AB* and *AD* events differ markedly between the two sets of studies. All of these differences make it difficult to generalize the conclusions derived from the literature on interference to associative recognition paradigms.

To the best of our knowledge, research on associative recognition has not yet explored the possibility of interference between distractor and target processing.[1] The aim of the following experiments was to fulfill this objective. To do so, a *yes–no* recognition test, in which half of the test pairs were intact and half rearranged, was repeatedly presented. Four categories of intact *AB* pairs were contrasted. What gave them their specific status was that none, one, or both of their constitutive members were also part of the rearranged pairs. In Category 1, none of the members of the pairs was part of the distractors. Using a *target/distractor* template, these pairs will be designated hereafter as *AB/XX*, with *X* standing for an item different from *A* and *B*, respectively. These pairs were used as the baseline. In Category 2, *A* was sometimes paired with an incorrect partner in the distractor items. These pairs will be called *AB/AX*. Taken jointly, these first two categories of test items provide the materials with which to investigate retroactive interference as defined in the traditional literature. Two other categories of test items were added. In Category 3, denoted *AB/XB*, *A* was never followed by an incorrect partner, but *B* followed a partner other than *A* in some distractor items. Finally, the two prior sources of mismatching were used to produce the last pair category

(Category 4), and thus, they will be designated as *AB/AX&XB*. There were two pairs in each of the four categories. The final composition of the test material is shown in Table 1.

Our primary question was the following: When people repeatedly perform a recognition test patterned as indicated in Table 1, do rearranged pairs influence the accuracy and the speed of responses to the intact pairs? Assuming a positive response, a second question was, Is this effect limited to or, at least, particularly marked for those of the distractors that are misconstrued as correct? If so, the effect should vary as a function of the rate of false alarms generated by the situation. To explore this issue, the amount of initial learning between arbitrary pairs was manipulated in Experiment 1, and Experiment 2 relied on associative knowledge previously established in a real-world context. Finally, a subsidiary question concerned the relative influence of the two possible sources of mismatches, which are referred to hereafter as *forward* (the effect of *AX* on *AB* remembering) and *backward* (the effect of *XB* on *AB* remembering), respectively. Are both kinds of effect detectable and, if so, are they cumulative?

EXPERIMENT 1

The first experiment involved the arbitrary pairing of visual Armenian letters with oral syllables. The participants were divided into three groups. Each group received a different amount of training before running a *yes–no* recognition test.

Table 1
Composition of the 16 *AB* Pairs Used in the *Yes–No* Recognition Tests in Experiments 1 and 2

	Number	*A*	*B*	Category
Targets				
(intact pairs)	1	1	1	*AB/XX*
	2	2	2	
	3	3	3	*AB/AX*
	4	4	4	
	5	5	5	*AB/XB*
	6	6	6	
	7	7	7	*AB/AX&XB*
	8	8	8	
Distractors				
(rearranged pairs)	9	3	7	
	10	3	8	
	11	4	7	
	12	4	8	
	13	7	5	
	14	7	6	
	15	8	5	
	16	8	6	

Note—*AB/XX*, no member of the pair is displayed elsewhere; *AB/AX*, the first member of the pair appears in distractors; *AB/XB*, the second member of the pair appears in distractors; *AB/AX&BX*, both members of the pair appear in distractors. For instance, *A3B3* (Pair 3) was an *AB/AX* pair, because *A3* was mismatched with *B7* (Pair 9) and *B8* (Pair 10) but *B3* never appeared in distractors.

Method

Participants. A total of 60 undergraduate students from the University of Bourgogne in Dijon, France, participated in the experiment in partial fulfillment of a course requirement. All the participants were native French speakers. The participants were randomly assigned to one of the three experimental groups (*n* = 20 for each group).

Materials. The eight *AB* pairs used in this experiment were composed of one Armenian letter and one syllable. The Armenian letters were selected from the Armenian alphabet in such a way as to be as different as possible from one another and, also, as different as possible from Latin letters, Arabic digits, and other familiar symbols. The black letters, approximately 20 mm high and 15 mm wide, were displayed on a white background on a computer screen. The syllables were {ba di ro tu la mi no vu}. They were synthesized using the MBROLA speech synthesizer (tcts.fpms.ac.be/synthesis/; Dutoit, Pagel, Pierret, Bataille, & Van Der Vrecken, 1996) with the FR2 diphone database. The resulting WAV files were played through headphones connected to a PC computer. The duration of each syllable was 350 msec.

Procedure. The participants were run individually in a sound-attenuated room. The experiment comprised a training phase and a test phase. Before the training phase, the participants were told that they would have to learn the name of a set of unknown letters shown on the computer screen. The to-be-learned material was then displayed. The Armenian letters and the syllables were arbitrarily mapped, with a different mapping for each participant. On each trial, an Armenian letter was displayed on the screen for 2 sec. The paired auditory syllable occurred with a stimulus onset asynchrony (SOA) of 1 sec. Each of the eight pairs occurred once in each block of trials. The pairs were pseudorandomly ordered for each participant within each block. The number of training blocks differed according to the groups. There were three groups of participants, who received 1 (Training 1), 3 (Training 3), and 10 (Training 10) blocks of trials, respectively.

After training, the participants were told that they would be presented again with a set of letter–sound pairs and that they would have to judge, for each of them, whether the pair was correct of not. They were asked to press the "C" key on the computer keyboard if the pair was intact and the "N" key if the pair was rearranged. The instructions emphasized both speed and accuracy. On each trial, a syllable was played first, followed by a letter on the screen, with an SOA of 350 msec. The letter was erased 1 sec after the participants had answered. In case of an erroneous response, the word ERROR! was printed just below the target item. Otherwise, no feedback was provided. There were 10 blocks of test trials, each block comprising 16 pairs, patterned as shown in Table 1. The pairs were given in random order, with the constraint that any given member of a pair was never repeated over four successive trials. The order of the pairs differed for each block and each participant.

Results

The following analyses deal successively with how well and how quickly the participants answered when confronted with the intact pairs. For all the ANOVAs, planned orthogonal contrasts were used to decompose the effect (if significant) of the category of pairs, in order to address successively three questions: (1) Did performance on baseline pairs (*AB/XX*) differ from performance on the other pairs (*AB/AX*, *AB/XB*, and *AB/AX&XB*), (2) did performance on associations with bidirectional mismatches differ from performance on associations with either backward or forward mismatches (*AB/AX&XB* vs. *AB/AX*, *AB/XB*), and (3) did performance differ according to directionality (*AB/AX* vs. *AB/XB*)? A final set of analyses was devoted to performances on distractors.

Errors on intact pairs. The overall rate of misses was 14.3%. An ANOVA was performed on the misses, with training level (1, 3, or 10 blocks of training) as a between-subjects factor and test ($N = 10$ repetitions of the recognition list) and pairs ($N = 4$) as within-subjects factors. There was a significant difference between the training levels [$F(2,57) = 8.64$, $p < .001$], with an error rate of 15.7%, 21%, and 6.25% for the groups trained with 1, 3, and 10 blocks of trials, respectively. Planned comparisons showed that the scores of the groups receiving 1 and 3 blocks of training did not differ significantly [$F(1,57) = 2.18$, $p = .14$], whereas these groups made significantly more errors than did the group receiving 10 training blocks [$F(1,57) = 15.09$, $p < .001$]. There was also a main effect of test [$F(9,513) = 15.82$, $p < .001$], due to the fact that the rate of errors went from 24.5% in Test 1 to 10.5% in Test 10, following a linear decreasing trend [$F(1,57) = 67.19$, $p < .001$]. Unsurprisingly, this decrement was more marked for the Training 1 group than for the groups receiving more extensive training, and this pattern elicited a significant test $\times$ training level interaction [$F(18, 513) = 2.10$, $p = .005$]. Note that the rate of misses was always far below the chance level (50% correct responses), even in Test 1 for the participants receiving only 1 or 3 blocks of training (32.5% and 25.2%, respectively; ts > 5.87, ps $< .001$).

Central to our concern, the main effect of pair was significant [$F(3,171) = 10.46$, $p < .001$]. This effect did not vary significantly across tests, as was revealed by the absence of a pair $\times$ test interaction [$F(27,1539) = 0.76$, $p = .81$], but differed according to training level [pair $\times$ training level interaction: $F(6,171) = 2.55$, $p = .022$]. The three-way interaction was not significant.

The effect of the category of pairs is shown for each training level in Figure 1. Despite the reliable training level $\times$ pair interaction, the effect of pair appears to follow roughly the same pattern for the three groups. The lowest rate of misses appeared in the *AB/XX* pairs, and the largest rate in the *AB/AX&XB* pairs, with the two intermediary categories falling in between. The only exception to this pattern (which certainly caused the training level $\times$ pair interaction) was the high rate of misses observed in the *AB/AX* pairs for the Training 1 group. Planned orthogonal comparisons showed a significant difference between the *AB/XX* pairs and the three other categories [$F(1,57) = 28.67$, $p < .001$]. Within the latter, the difference between the pairs with both forward and backward distractors (*AB/AX&XB*) and the other categories of pairs (*AB/AX* and *AB/XB*) was marginally significant [$F(1,57) = 3.55$, $p = .064$], whereas *AB/AX* and *AB/XB* pairs did not reliably differ from each other [$F(1,57) = 1.069$, $p = .305$].

Reaction times for intact pairs. Reaction times (RTs) were measured as the interval between the moment a letter appeared on the screen and the moment the participant pressed a key. RTs above 3,500 msec were removed from the analyses. They represented 0.125% of all the values.[2] An ANOVA was performed on the RTs for hits, with training level (1, 3, and 10) as a between-subjects factor and test ($N = 10$) and pairs ($N = 4$) as within-subjects factors. The training level generated no main effect and did not interact with the other factors [$F_{max}(2,57) = 1.14$, $p_{min} = .304$]. RTs depended reliably on both test [$F(9,513) = 41.35$, $p < .001$] and pair [$F(3,171) = 12.99$, $p < .001$], and there was a significant interaction between the two factors [$F(27,1539) = 1.87$, $p < .001$].

Figure 2 shows the mean RTs on correct responses for the four categories of pairs as a function of test repetitions, after averaging on the training levels. RTs improved across the tests for all the conditions, but at a different rate as a function of pairs. Planned orthogonal comparisons showed that the *AB/XX* pairs elicited shorter RTs than did the pairs with related distractors [$F(1,57) = 32.84$, $p < .001$]. Within the latter categories, the pairs with either forward or backward mismatches elicited shorter RTs than
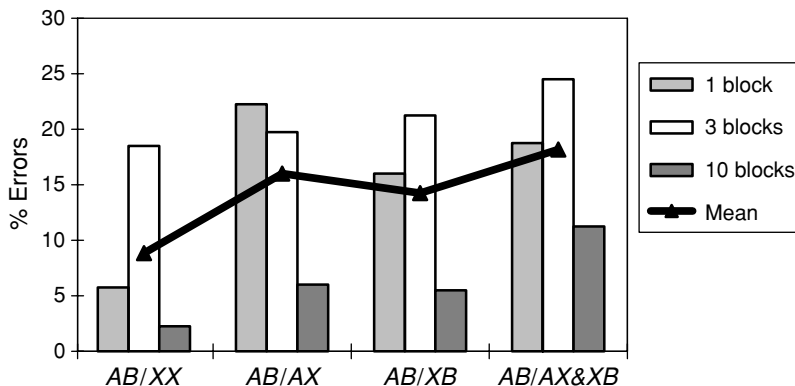

Figure 1. Rates of errors for intact pairs (misses) as a function of (1) the number of prior training blocks and (2) whether the members of the pair were also displayed in distractors (*AB/XX*, no member of the pair was displayed elsewhere; *AB/AX*, the first member of the pair appeared in distractors; *AB/XB*, the second member of the pair appeared in the distractors; *AB/AX&BX*, both members of the pair appeared in distractors).
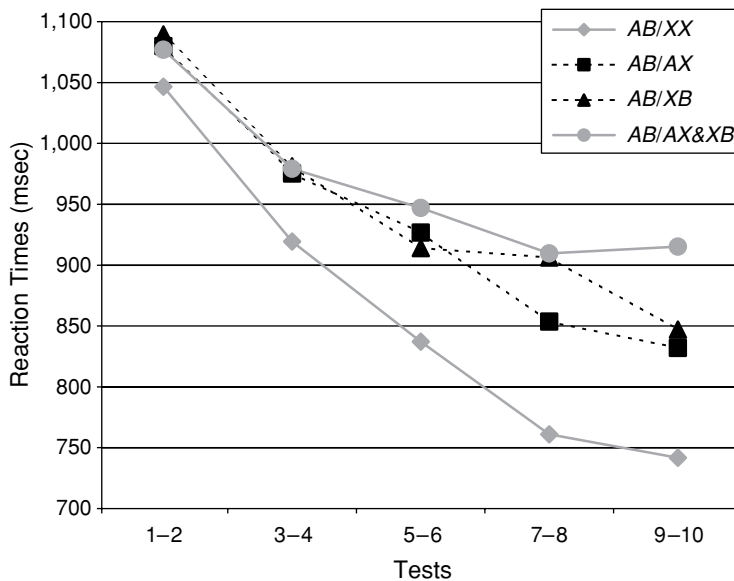
Figure 2. Reaction times for correct recognition of intact pairs (hits) as a function of whether the members of the pair were also displayed in the distractors (see the legend of Figure 1 for details).

did the pairs with bidirectional mismatches (*AB/AX&XB*), but the effect was only marginally significant [$F(1,57) = 3.25, p = .077$], and there was no reliable difference between *AB/AX* and *AB/XB* [$F(1,57) = 0.55, p = .46$]. In order to analyze the last two effects more precisely, the same contrasts were computed after restricting the data to the last two blocks of testing, where the effect of testing was assumed to be maximal. The difference between *AB/AX&XB* and the other two categories with related distractors was now significant [$F(1,57) = 10.42, p = .002$], but *AB/AX* and *AB/XB* pairs still did not differ [$F(1,57) = 0.26, p = .61$].

Performance on rearranged pairs. The mean rate of errors on distractors (false alarms) was 14.25%. The scores differed between groups, with a false alarm rate of 19% for the Training 1 group, 18.6% for the Training 3 group, and 5.1% for the Training 10 group [$F(2,57) = 7.93, p < .001$]. Planned comparisons showed that the Training 10 group made a lower rate of false alarms than did the two other groups [$F(1,57) = 15.86, p < .001$], which did not differ from each other [$F(1,57) = 0.01, p = .92$]. The number of false alarms significantly decreased across training, passing from 23.7% to 11.2% from Test 1 to Test 10 [$F(9,513) = 17.61, p < .001$; linear trend: $F(1,57) = 78.70, p < .001$]. This effect differed as a function of training level, as is shown by the significant test × training level interaction [$F(18,513) = 8.05, p < .001$]. This interaction was due to the fact that the participants in the Training 10 group started with a lower rate of false alarms, which limited the possibility of further decrement. The mean RTs for correct responses (correct rejections) also varied as a function of test [$F(9,513) = 33.31, p < .001$], following a negative linear trend [$F(1,57) = $

107.98, $p < .001$]. However, by contrast with the analysis on errors, neither the main effect of training level nor the training level × test interaction was significant ($Fs < 1$).

To sum up, the repeated presentation of a recognition test had a clear beneficial effect on both accuracy and speed of responding, and this was true for the responses to intact and rearranged pairs. However, the rate of improvement for intact pairs differed according to whether the distractors involved none, one, or two of their constituent members. Overall, a relative impairment in performance on intact pairs was observed when the distractors involved one or the other of the members of the pairs, with partial evidence for a cumulative effect when the two members of the pairs were shared by the distractors. These effects did not vary reliably as a function of whether the participants had received 1, 3, or 10 training blocks, although there were sizable differences in the rate of false alarms between the latter group and the two former ones.

EXPERIMENT 2

The results of Experiment 1 raised two questions. First, the fact that the deleterious impact of distractors was unrelated to the amount of initial training suggests that it could be independent from whether the distractors are construed as correct or not. It is also possible, however, that the variation that we introduced in the amount of initial training was not large enough to reveal an effect. The main objective of Experiment 2 was to replicate Experiment 1 with overlearned associations. A second question stemmed from the observation that backward mismatches (*XB*) were as deleterious as forward mismatches (*AX*). A somewhat trivial interpretation of this effect could be that

the participants did not process the association of an Armenian letter and an oral syllable as an oriented relation. Indeed, the relationship between a grapheme and a phoneme is clearly bidirectional in literacy activities, depending on whether reading or writing is involved. Moreover, whereas the syllable was played during the presentation of the visual letter in the training phase, the syllable preceded the presentation of the visual letter in the recognition test in Experiment 1, and this may have contributed to an overshadowing of the direction of the relation.

In order to address these two issues, we used small multiplication facts as materials in Experiment 2. They are overlearned through both school instruction and everyday practice, and the processing of multiplications is clearly oriented (although the equation $48 = 6*8$ is mathematically correct, it is taught and generally used in the reverse order, for obvious reasons).

Method

Participants. A total of 20 undergraduate students from the same pool as that in the prior experiment participated in partial fulfillment of a course requirement. An additional participant who made a very large rate of errors (more than seven standard deviations above the mean of the sample) was removed and replaced. All the participants were native French speakers.

Materials. The 16 *AB* pairs used in the test were composed of two single-digit operands as their *A* member (e.g., 6×8) and a number, which was either the correct answer (48) or an incorrect one (e.g., 56), as their *B* member. The problems were displayed with an "=" sign between the *A* and the *B* members, the resulting string (e.g., $6 \times 8 = 48$) filling a rectangle of approximately $40*8$ mm on the computer screen.

As in Experiment 1, the pairs were patterned as shown in Table 1, in order to generate four categories of correct pairs (*AB/XX*, *AB/AX*, *AB/XB*, and *AB/AX&XB*). In Experiment 1, the assignment of a specific pair to an experimental category was randomized on an individual basis. Such a randomization did not appear desirable with arithmetic facts, because some distractors would have looked trivially false, hence making the task strange or meaningless for the participants. In Experiment 2, the distractors were built in order to look as "plausible" as possible, through the use of various ploys (e.g., the false answer was often part of the multiplication table of at least one of the operands and/or shared one digit with the correct answer).

However, in order to avoid any confounding between a specific multiplication fact and an experimental condition, four lists were prepared in such a way that over all the lists, each multiplication fact was ascribed once to each of the four experimental categories (see the Appendix).

Procedure. The participants were run individually in a sound-attenuated room. Because adult people presumably know multiplication facts, there was no training phase. The participants were told that they would have to check small multiplication problems shown on the computer screen. They were asked to press the "C" key on the computer keyboard if the operation was correct and the "N" key if the operation was incorrect, and the instructions emphasized both speed and accuracy. On each trial, the whole string was displayed on the screen at once, and remained until the participants answered. In case of an erroneous response, the word ERROR! was printed at the bottom of the screen for 2 sec. By default, the absence of feedback meant that the participants' response was correct. There were 10 blocks of trials, with each block comprising 16 items. The participants were randomly assigned to one of the four test sets shown in the Appendix, until there were 5 participants for each set. The items were presented in random order, with the constraint that any given string, whether the problem (*A*) or its answer (*B*), was never repeated over four successive trials. The order of the items differed for each block and each participant.

Results

Errors on correct operations. Overall, the participants made 7.75% "incorrect" responses when they were shown the right multiplication facts (misses). An ANOVA was performed on these errors, with test ($N = 10$) and pair ($N = 4$) as within-subjects factors. The main effect of test [$F(9,171) = 0.91$, $p = .517$] and the test $\times$ pair interaction [$F(27,513) = 0.95$, $p = .543$] were not significant. However, the category of pairs generated reliable differences [$F(3,57) = 5.67$, $p = .002$]. The effect of pair is shown in Figure 3. The difference between the extreme categories was striking, with the number of misses being almost four times larger for *AB/AX&XB* pairs than for *AB/XX* pairs. The two intermediary categories of pairs were in between. Planned orthogonal comparisons showed a significant difference between the baseline and the three other categories [$F(1,19) = 13.72$, $p = .002$]. However, the difference between the pairs with both forward and
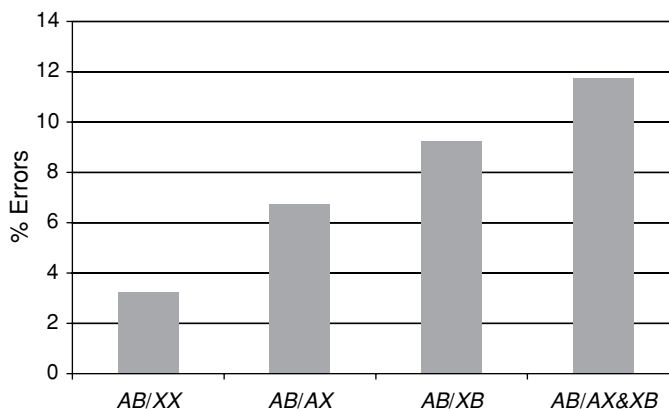

Figure 3. Rate of errors on right multiplication facts as a function of whether the members of the pair were also displayed in the distractors (see the legend of Figure 1 for details).

backward mismatches (*AB/AX&XB*) and the other two (*AB/AX* or *AB/XB*) did not reach significance [$F(1,19) = 2.61, p = .122$], and the latter did not significantly differ from each other [$F(1,19) = 2.21, p = .154$].

Reaction times for correct operations. As in Experiment 1, the analysis on RTs was performed on the correct *yes* responses (hits). The values above 5,000 msec, which made up 0.37% of the responses, were removed. An ANOVA was performed on RTs with test ($N = 10$) and pair ($N = 4$) as within-subjects factors. There was a significant effect of test [$F(9,171) = 9.21, p < .001$] and a marginally significant effect of pair [$F(3,57) = 2.60, p = .061$]. The test $\times$ pair interaction failed to reach significance [$F(27,513) = 1.09, p = .346$].

Figure 4 shows mean RTs as a function of tests and pairs. RTs followed a decreasing trend for all pairs. However, RTs for *AB/XX* pairs were shorter than those for the other three categories [$F(1,19) = 7.56, p = .013$], whereas the latter did not differ from one another [$F(2.38) = 0.09, p = .91$].

Performance on false operations. The participants made 11.19% *yes* responses on erroneous operations (false alarms). This value differed between tests in a marginally significant way [$F(9,171) = 1.81, p = .070$], but no systematic trend emerged, with a nearly null slope of the adjustment line [$F(1,19) = 0.01, p = .97$]. By contrast, the RTs for the correct rejection of the false operations differed between tests [$F(9,171) = 11.19, p < .001$], with a sizable decrement from Test 1 to Test 10 [linear trend: $F(1,19) = 17.02, p < .001$].

Experiment 2 revealed that the detrimental impact of distractors on accuracy and speed of responses to correct items persists when correct items have been overlearned in the first place. However, examining the data could lead one to argue that the students' mastery of multiplication

facts may have been limited. For instance, the mean RTs to multiplication facts in Experiment 2 were *longer* than those for arbitrary pairings in Experiment 1. It should be unwarranted to infer too much from this comparison, because the procedures of the two experiments were not identical.[3] Nevertheless, it must be acknowledged that there was no evidence indicating that the processing of multiplication facts was fully automatized by the participants. Our only, hardly disputable claim is that prior training on multiplication facts was more extensive than training on the arbitrary symbol–syllable pairs in Experiment 1.

GENERAL DISCUSSION

A Strong Interference Effect

Our main question was the following: Do distractors influence the accuracy and the speed of responses to targets in associative recognition tasks when the former share common elements with the latter? Our data unambiguously support the conclusion that rearranged pairs have harmful consequences on the memory for correct pairs. It is worth emphasizing that the detrimental effect of distractors was *relative*. In all the conditions, there was a clear improvement in performance across the repetitions of the test for both speed and accuracy measures. However, the speed and accuracy of responses to the intact pairs sharing common elements with the distractors improved to a lesser extent than did the corresponding measures for the intact pairs that did not have any element in common with the distractors. The effect was significant in the two experiments for both errors and RTs. The impact of distractors on RTs significantly increased across tests in Experiment 1, and the same trend appeared in Experiment 2, although it failed to reach significance. A visual inspec-
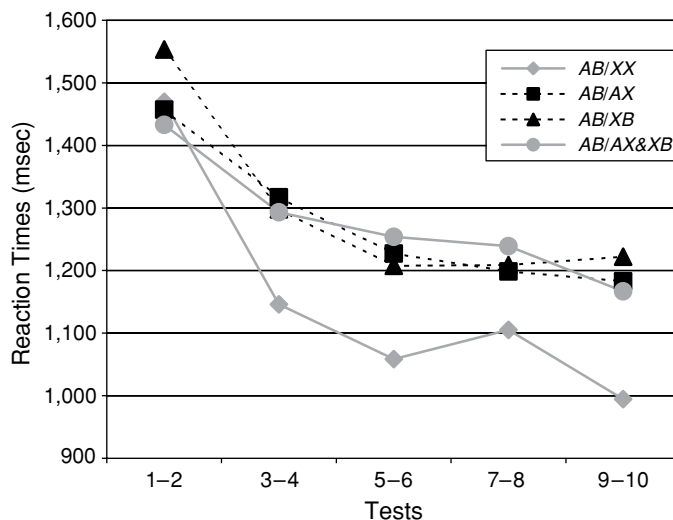

Figure 4. Reaction times for correct responses to right multiplication facts as a function of whether the members of the multiplication facts were also displayed in the distractors (see the legend of Figure 1 for details).

tion of Figures 2 and 4 shows that in both experiments, the impact of distractors was quite sizable from Tests 3 and 4 upward.

In order to assess the size of the harmful effect of distractors, we compared the scores in the baseline condition (*AB/XX*) and the mean scores in the other three conditions, using Cohen's *d* (e.g., Cohen, 1988) as an index for effect size. Regarding first the errors, the rate of misses jumped from 8.83% (*SE* = 1.75) in the baseline condition to 16.14% (*SE* = 1.77) in the interfering conditions in Experiment 1. The corresponding values in Experiment 2 were 3.25% (*SE* = 0.87) and 9.25% (*SE* = 1.49). Cohen's *d*s on these error scores were 0.536 and 0.634 for Experiments 1 and 2, respectively. Regarding RTs on hits, the difference in RTs, pooled over the three levels of training and the 10 successive tests, was 87.7 msec (*SE* = 15.09) for Experiment 1. For Experiment 2, the difference was still more impressive, reaching 129.05 msec (*SE* = 27.09). Cohen's *d*s on RTs were 0.507 and 1.217 for Experiments 1 and 2, respectively. Considering Cohen's effect size conventions (*d* = 0.2, small; *d* = 0.5, medium; *d* = 0.8, large), these effects fell roughly in the medium range for the arbitrary pairs in Experiment 1 and in the large range for multiplication facts in Experiment 2.

Is the Effect Due to Those Distractors That Were Endorsed as Correct?

If the detrimental effect of distractors was due to those distractors that were incorrectly endorsed as correct, the amount of interference would depend on the proportion of false alarms that the situation elicited. Our results provided no support for this prediction. First, the comparison between the two experiments suggests some advantage for the opposite conclusion. Indeed, although Experiment 1 dealt with arbitrary associations learned during an experimental training session (with a mean false alarm rate of 14.23%) and Experiment 2 with overlearned multiplication facts (with a mean false alarm rate of 11.19%), the size of the effect was larger in the second case. Of course, this comparison is questionable, because the two experiments differed with regard to many other factors. For instance, the second experiment involved numerical data, which are known to be highly susceptible to interference (e.g., Zbrodoff & Logan, 1986).

Nevertheless, within-experiment evidence converges toward the same conclusion. The amount of training was manipulated in Experiment 1, and no consistent difference in the effect of mismatching was observed between participants trained over 1, 3, or 10 blocks of trials, although the rate of false alarms significantly differed between the first two groups (19% and 18.6%, respectively) and the last one (5.1%). In order to explore further the role of false alarms, we ran an additional analysis on RTs, contrasting, within each group, the half of the participants who made the lowest rate of false alarms with the other half. There was a striking difference in the rate of false alarms produced by the "good" and the "poor" learners (5.56% vs. 22.87%, respectively). However, the effect of distractors was nearly identical. The difference between

the baseline condition (*AB/XX*) and the other three conditions was 88 msec (857 vs. 945 msec) for good learners and 87 msec (865 vs. 952 msec) for poor learners. A final analysis was restricted to the good learners from the group that had received the longest training. Although those participants made virtually no false alarms (0.62%), the difference between the baseline condition and the three interfering conditions was unchanged [87 msec; *F*(1,54) = 5.44, *p* = .023].

A similar analysis was performed for Experiment 2. Again, there was a large difference between the false alarm rate for the two groups of participants (3.75% vs. 19.12%). However, the detrimental effect of distractors was, in fact, *stronger* for good than for poor calculators. As assessed by the difference between the baseline condition and the interfering conditions, the effect was 200 msec (1,154 vs. 1,354 msec) for good calculators and 58 msec (1,155 vs. 1,213 msec) for poor calculators (note that when these data were entered into an ANOVA designed as above, but with a new between-group factor contrasting good and poor calculators, the interaction between this new factor and the status of the pairs did not reach significance [*F*(3,54) = 1.08, *p* = .365]. In summary, although the occurrence of false alarms attests that some distractors were processed as if they were correct, it appears quite unlikely that this factor would play a major role in interference generation. Those results provide evidence for the fact that blatantly incorrect information may impede previously established knowledge to a substantial extent.

About Forward and Backward Rearrangements

Finally, our experimental design provided the opportunity to investigate the effect of two categories of distractors, which we called *forward* and *backward*, respectively, as a function of the directionality of the target/distractor mismatches. Is the influence of backward mismatches similar to the influence of forward mismatches, and if so, are the two influences cumulative? Regarding the first question, the response is clearly positive: The influence of backward mismatches was indistinguishable from the influence of forward mismatches for both accuracy and RTs. This was true in Experiment 1, although the interpretation of this result can be challenged with the argument that the direction of the association was ambiguous. However, the same effect occurred in Experiment 2, even though a multiplication fact is clearly oriented from a problem toward its answer. Whether the two effects were cumulative is more debatable. There was a trend for a cumulative effect with arbitrary associations, which was marginally significant for both errors and RTs. With multiplication facts, we observed only a descriptive, nonsignificant trend for a cumulative effect, which was limited to accuracy measures.

Implications

In the context of the literature on associative recognition, our design included some unusual aspects, such as the repeated presentation of a small number of selected distractors. Presumably, the amount of interference elic-

ited in more typical conditions should be of lesser extent. This is an empirical issue that needs to be settled in further studies before assessing whether the occurrence of interference might have a significant impact for the current research on associative recognition.

However, irrespective of the response to this question, our results have major implications in other research contexts. As a case in point, most of the pedagogical devices for computer-aided instruction involve multiple-choice procedures. Because the correct response is displayed within a set of distractors, the latter must necessarily be encoded for the correct response to be selected. Marsh and Roediger (2002) have shown that the distractors presented in a multiple-choice test were produced more frequently in a subsequent free recall test of general knowledge. However, it is unclear whether participants construed these distractors as false while performing the multiple-choice test. Our study suggests that the processing of distractors may impede the knowledge of correct associations, even if prior training has been extensive enough to prevent the choice of the erroneous option. This result is reminiscent of the conclusions from earlier studies on spelling. For instance, Brown (1988) explored the harmful effect of reading defective writings on subsequent spellings. He presented erroneous spellings of words to participants who were asked to rate how closely those misspellings resembled the correct spellings. The subsequent recognition of the correct spelling was impaired by the prior exposure to the incorrect alternatives (see also Jacoby & Hollingshead, 1990; Nisbet, 1939). Thus, interference occurred even though the participants were not instructed to learn the misspellings and were fully aware of their incorrectness. Together with studies on spelling, our report provides a further incitement to proscribe any pedagogical methods that could favor exposure to false association and, in general, to other forms of erroneous information, even when the intention is to test presumably well-consolidated knowledge.[4]

REFERENCES

ANDERSON, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory & Language*, **49**, 415-445.

ANDERSON, M. C., BJORK, E. L., & BJORK, R. A. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1063-1087.

ANDERSON, M. C., & NEELY, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory: Handbook of perception and cognition* (2nd ed., pp. 237-313). San Diego: Academic Press.

BEGG, I. M., ANAS, A., & FARINACCI, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, **121**, 446-458.

BROWN, A. S. (1988). Encountering misspellings and spelling performance: Why wrong isn't right. *Journal of Educational Psychology*, **80**, 488-494.

CLEARY, A. M., CURRAN, T., & GREENE, R. L. (2001). Memory for detail in item versus associative recognition. *Memory & Cognition*, **29**, 413-423.

COHEN, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

DELPRATO, D. J. (2005). Retroactive interference as a function of degree of interpolated study without overt retrieval practice. *Psychonomic Bulletin & Review*, **12**, 345-349.

DUTOIT, T., PAGEL, N., PIERRET, F., BATAILLE, O., & VAN DER VRECKEN, O. (1996). The MBROLA Project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. In H. T. Bunell & W. Isardi (Eds.), *Proceedings of the Fourth International Conference on Spoken Language Processing* (pp. 1393-1396). Wilmington, DE: Alfred I. du Pont Institute.

GERRIG, R. J., & PRENTICE, D. A. (1991). The representation of fictional information. *Psychological Science*, **5**, 336-340.

HASHER, L., GOLDSTEIN, D., & TOPPINO, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning & Verbal Behavior*, **16**, 107-112.

HOCKLEY, W. E. (1991). Recognition memory for item and associative information: A comparison of forgetting rates. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 227-248). Hillsdale, NJ: Erlbaum.

JACOBY, L. L., & HOLLINGSHEAD, A. (1990). Reading student essays may be hazardous to your spelling: Effects of reading incorrectly and correctly spelled words. *Canadian Journal of Psychology*, **44**, 345-358.

LIGHT, L. L., PATTERSON, M. M., CHUNG, C., & HEALY, M. R. (2004). Effects of repetition and response deadline on associative recognition in young and older adults. *Memory & Cognition*, **32**, 1182-1193.

MARSH, E. J., & ROEDIGER, H. L., III (2002, November). *The effects of taking a multiple-choice test of later production of facts*. Poster presented at the 44th Annual Meeting of the Psychonomic Society, St. Louis.

NISBET, S. D. (1939). Non-dictated spelling tests. *British Journal of Educational Psychology*, **9**, 29-44.

TOPPINO, T. C., & BROCHIN, H. A. (1989). Learning from tests: The case of true–false examination. *Journal of Educational Research*, **83**, 119-124.

WESTERMAN, D. L. (2001). The role of familiarity in item recognition, associative recognition, and plurality recognition on self-paced and speeded tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 723-732.

ZBRODOFF, N. J., & LOGAN, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General*, **115**, 118-130.

NOTES

1. Prior studies have shown that on recognition or multiple-choice tests, participants rate distractors as being more correct on their successive occurrences. This effect is usually referred to as the *negative suggestion effect* (e.g., Toppino & Brochin, 1989). It may be construed as an instance of a more general phenomenon—namely, that as a consequence of its repetition, false information tends to be perceived as increasingly true (this phenomenon has been coined the *illusory truth effect*; see, e.g., Begg, Anas, & Farinacci, 1992; Gerrig & Prentice, 1991; Hasher, Goldstein, & Toppino, 1977). Note, however, that the focus of this literature is on the fate of distractors (or more generally, incorrect display), whereas our concern is about the fate of correct knowledge.

2. In Experiments 1 and 2, the threshold for outlier exclusion was selected after a careful inspection of the RT distribution, in order to remove only the values that were clearly apart from the others. In any case, the removed values represented a very small proportion of the total sample.

3. In Experiment 1, the auditory syllables were played (and presumably coded) before the Armenian letters, from which the RT was measured. In Experiment 2, the pairs were displayed at once, and hence, RTs comprised the time needed to read the whole strings, which were composed of several, nonredundant digits and symbols.

4. Although the present study was primarily intended to address empirical questions, our results have also some implications for the theories of retroactive interference. For instance, in the view recently proposed by M. C. Anderson (e.g., Anderson, 2003; Anderson, Bjork, & Bjork,

1994), forgetting of the pairs initially learned is thought of as a controlled response that occurs while people are trying to learn new pairings. According to their view, the unwanted retrieval of the initial pairs would trigger executive control mechanisms oriented toward the active suppression of those memories. A logical implication of this theory is that interference should not occur when the potentially interfering information is conveyed by the distractors in a *yes/no* recognition test. Indeed, participants do not attempt to learn the test items, and hence, they have no reason to suppress previously established knowledge. On the contrary, the suppression of previously established knowledge can only have a devastating effect, since it entails the loss of the participants' ability to discriminate the targets from the distractors in recognition tests. Our observation that retroactive interference still occurs in those conditions suggests that controlled suppression cannot be construed as the unique and, perhaps, even the main cause of retroactive interference (see also Delprato, 2005).

APPENDIX

	Number	Set 1 A	Set 1 B	Set 2 A	Set 2 B	Set 3 A	Set 3 B	Set 4 A	Set 4 B	Category
Intact	1	6×8	48	4×9	36	7×4	28	9×5	45	*AB/XX*
	2	8×4	32	9×6	54	3×8	24	6×7	42	
	3	7×4	28	9×5	45	6×8	48	4×9	36	*AB/AX*
	4	3×8	24	6×7	42	8×4	32	9×6	54	
	5	9×5	45	6×8	48	4×9	36	7×4	28	*AB/XB*
	6	6×7	42	8×4	32	9×6	54	3×8	24	
	7	4×9	36	7×4	28	9×5	45	6×8	48	*AB/AX&XB*
	8	9×6	54	3×8	24	6×7	42	8×4	32	
Rearranged	9	7×4	36	9×5	28	6×8	45	4×9	48	
	10	7×4	54	9×5	24	6×8	42	4×9	32	
	11	3×8	36	6×7	28	8×4	45	9×6	48	
	12	3×8	54	6×7	24	8×4	42	9×6	32	
	13	4×9	45	7×4	48	9×5	36	6×8	28	
	14	4×9	42	7×4	32	9×5	54	6×8	24	
	15	9×6	45	3×8	48	6×7	36	8×4	28	
	16	9×6	42	3×8	32	6×7	54	8×4	24	