CRITERION LEARNING AND ASSOCIATIVE ASYMMETRY: INVESTIGATING
THE ROLE OF MEDIATORS

A dissertation submitted to
Kent State University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

by
Kalif E. Vaughn

March 2014
Dissertation written by
Kalif E. Vaughn
B.A., Purdue University, 2008
M.A., Kent State University, 2013
Ph.D., Kent State University, 2014

Approved by
Katherine A. Rawson          Chair, Doctoral Dissertation Committee
Maria S. Zaragoza            Members, Doctoral Dissertation Committee
Jeffrey A. Ciesla
Brad Morris
Brian Huot

Accepted by
Maria S. Zaragoza            Chair, Department of Psychology
James L. Blank               Dean, College of Arts and Sciences
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ iv
LIST OF TABLES ........................................................................................................ vi
INTRODUCTION ......................................................................................................... 1
Measuring Associative Symmetry .............................................................................. 3
Evidence for Asymmetric Effects of Criterion Level on Associative Memory .......... 5
Overview of Current Research ............................................................................... 11
EXPERIMENT 1 ...................................................................................................... 18
Method .................................................................................................................. 18
Participants and Design ....................................................................................... 18
Materials ................................................................................................................. 18
Procedure ............................................................................................................... 19
Results and Discussion ......................................................................................... 21
EXPERIMENT 2 ...................................................................................................... 25
Method .................................................................................................................. 27
Participants and Design ....................................................................................... 27
Materials and Procedure .................................................................................... 28
Results and Discussion ......................................................................................... 29
Summary ................................................................................................................. 42
EXPERIMENT 3 ...................................................................................................... 44
Method .................................................................................................................. 45
Participants and Design ....................................................................................... 45
Materials ................................................................................................................. 46
Procedure ............................................................................................................... 47
Results and Discussion ......................................................................................... 49
Summary ................................................................................................................. 56
GENERAL DISCUSSION ...................................................................................... 58
LIST OF FIGURES

1 Mean final test performance as a function of criterion level and final test format in Experiment 1 (left column) and Experiment 2 (right column) from Vaughn and Rawson (resubmitted). Performance is reported in percentages, and error bars report standard error of the mean.

2 Mean final cued-recall performance as a function of criterion level, type of cued recall test, and pre-exposure condition in Experiment 1. Performance is reported in percentages, and error bars report standard error of the mean.

3 Mean final recognition performance as a function of criterion level, type of recognition test, and pre-exposure condition in Experiment 1. Performance is reported in percentages, and error bars report standard error of the mean.

4 Mean final cued-recall performance as a function of criterion level and type of cued recall test in Experiment 2. Performance is reported in percentages, and error bars report standard error of the mean.

5 Mean final recognition performance as a function of criterion level and type of recognition test in Experiment 2. Performance is reported in percentages, and error bars report standard error of the mean.

6 Mean final cued-recall performance for items in which participants reported using a verbal mediator during practice (left panel), a visual mediator during practice (middle panel), or some other kind of strategy (including no strategy or repetition)
during practice (right panel), represented as a function of criterion level and type of cued recall test in Experiment 2. Performance is reported in percentages, and error bars report standard error of the mean.

7 Mean final recognition performance for items in which participants reported using a verbal mediator during practice (left panel), a visual mediator during practice (middle panel), or some other kind of strategy (including no strategy or repetition) during practice (right panel), represented as a function of criterion level and type of recognition test in Experiment 2. Performance is reported in percentages, and error bars report standard error of the mean.

8 Mean final cued-recall performance as a function of mediator group and type of cued recall test in Experiment 3. Performance is reported in percentages, and error bars report standard error of the mean.

9 Mean final recognition performance as a function of mediator group and type of cued recall test in Experiment 3. Performance is reported in percentages, and error bars report standard error of the mean.

10 Mean final cued-recall performance as a function of mediator strategy and type of cued recall test in Experiment 3. Performance is reported in percentages, and error bars report standard error of the mean.

11 Mean final recognition performance as a function of mediator strategy and recognition stimulus in Experiment 3. Performance is reported in percentages, and error bars report standard error of the mean.
LIST OF TABLES

1. Mean percentage of strategy use reported during the retrospective strategy reports in Experiments 2 and 3.
INTRODUCTION

Research has demonstrated that testing benefits subsequent memory (e.g., Rawson & Dunlosky, 2011). Numerous factors have been shown to moderate the benefits of testing, including retrieval success (e.g., correct vs. incorrect recall), the amount of testing (e.g., 1 vs. 2 test trials), and the test format implemented (e.g., recall vs. recognition). Perhaps surprisingly, a minimal amount of research has manipulated the amount of retrieval success during practice (hereafter referred to as criterion level). Nonetheless, a clear pattern has emerged from the limited research investigating criterion level: As criterion level increases during practice, subsequent final test performance also increases.

For instance, Pyc and Rawson (2009) had participants study and retrieve 70 Swahili-English word pairs via cued recall during practice. During a cued recall trial, the Swahili word would be presented as the cue (wingu) and participants were instructed to retrieve the correct English translation (referred to as the target word; in this case, cloud). If a participant correctly recalled the item, the computer would add one correct recall to that specific item. Items were pre-assigned to a specific criterion level (ranging from 1 to 10 correct recalls). If the item had reached the pre-assigned criterion level, then the item was removed from the remaining to-be-learned items and received no additional test practice. If the item had not yet reached their pre-assigned criterion level, then the item was placed at the end of the stack of the remaining to-be-learned items to receive
additional test practice. If the participant did not correctly recall the item, the computer would present the Swahili-English word pair for a restudy period (e.g., *wingu-cloud*). Items not correctly recalled were then placed at the end of the remaining stack of to-be-learned word pairs and were presented for an additional test trial later. The experiment ended when participants had reached the pre-assigned criterion level for each item. Overall, the results from Pyc and Rawson (2009) demonstrated a clear pattern: Final cued recall performance increased as a function of initial criterion level during practice. The basic pattern of increased final test performance given higher initial criterion level has been replicated (e.g., Vaughn & Rawson, 2011).

Recent research on criterion level effects has extended to explore what aspects of memory are enhanced via criterion level. For instance, Vaughn and Rawson (2011) had participants learn Lithuanian-English word pairs. After the word pairs had been presented for an initial study trial, participants completed practice cued recall tests until they correctly retrieved the word pairs a pre-assigned number of times (from one to five). Importantly, the practice cued-recall tests always occurred in the forward direction (i.e., participants retrieved the English translation given the Lithuanian word). The final test phase assessed forward cued recall (the same test format used during practice) and backward cued recall (i.e., participants retrieved the Lithuanian translation given the English word). Despite practicing only forward cued recall, performance increased as criterion level increased for both forward and backward cued recall. Although performance increased in both cued recall directions, performance gains were much greater in the forward versus backward direction (35 vs. 12%, respectively). The discrepancy in performance gains between forward and backward cued recall suggests that criterion level produces asymmetric effects on associative memory.
Measuring Associative Symmetry

Although Vaughn and Rawson (2011) provides suggestive evidence that criterion level produces asymmetric effects on associative memory, the study was not originally designed to evaluate associative asymmetry and thus there are methodological considerations that preclude a strong conclusion regarding associative asymmetry. In a recent review on associative symmetry, Kahana (2002) highlights two methodological issues that need to be addressed in order to evaluate associative symmetry. The first methodological issue concerns the extent to which the cue and target words have a similar number of pre-existing associates. As Kahana (2002) outlines, words with a higher number of pre-existing associates may produce more interference during cued recall, decreasing the chance of successfully retrieving the lesser associated word within the word pair (e.g., if the number of pre-existing associates is greater for a target word than for a cue word, the cue word would be more difficult to retrieve during backward cued recall due to more interference from other associates of the target). The second methodological issue concerns accounting for differences in item accessibility, because items with higher accessibility are more easily retrieved during cued recall (e.g., higher levels of item accessibility for target words versus cue words would benefit forward versus backward cued recall).

There are several ways to minimize concerns over differences in pre-existing associations, including drawing items from the same class of stimuli (e.g., using noun-noun word pairs) and then counterbalancing assignment of the items to be cues and targets (e.g., A-B and B-A). Many prior studies investigating associative symmetry violate one or both of these important prerequisites (e.g., Gallup & Wollen, 1968; Levy & Nevill, 1974; Lockhart, 1969; Bartling & Thompson, 1977) making it difficult to interpret their results with respect to associative symmetry. To foreshadow, in the current experiments, I minimized these concerns by using
unrelated noun-noun word pairs and by counterbalancing assignment of words to serve as either the cue or target during practice (e.g., “queen-journal” and “journal-queen”).

Additionally, there are several ways to minimize concerns over differences in item accessibility, such as including measures to assess both cue and target accessibility and/or implementing a procedure to equate cue and target accessibility (e.g., additional exposure to cue or target words, depending on which items have lower accessibility). Many prior studies investigating associative symmetry did not measure cue or target accessibility and/or made no attempt to equate cue or target accessibility (e.g., Wollen, 1968; Wollen & Allison, 1968; Newman & Campbell, 1971; Giurantano, 1972; Caplan, Graholt, & McIntosh, 2006). For example, many studies involved participants practicing in a forward direction (e.g., A-?) followed by subsequently demonstrating superior forward versus backward cued recall (e.g., Lockhart, 1969; Gallup & Wollen, 1968; Giurintano, 1972). However, no measures of target or cue accessibility were included, and item accessibility differences likely arose due to differential strengthening of items during practice. Forward cued recall practice presumably enhances memory strength for the target word in addition to strengthening the cue-target association. If so, the demonstrated advantage of forward versus backward cued recall may have reflected greater target accessibility rather than associative asymmetry. To foreshadow, I minimized these concerns by including measures of cue and target accessibility and by pre-exposing cues to equate cue and target accessibility in all experiments.

In sum, much of the prior research on associative symmetry is not readily interpretable due to one or more of these methodological limitations. Additionally, none of this earlier research investigating associative symmetry involved a manipulation of criterion level. Outcomes of the only criterion level study that included relevant measures (Vaughn & Rawson,
2011, described above) suggested that the effects of criterion level on associative memory are asymmetric. However, the study by Vaughn and Rawson (2011) was not originally designed to evaluate associative symmetry, and not surprisingly, the materials used violated the aforementioned methodological conditions necessary to evaluate associative symmetry. For example, the number of pre-existing associates differed between cue and target words, as Lithuanian versus English words have a lower number of pre-existing associates for native English speakers. Furthermore, performance on free recall and recognition measures suggested poorer item accessibility for cue words than for target words. Given these methodological limitations, the results of Vaughn and Rawson (2011) only provide suggestive evidence at best for associative asymmetry. Thus, the extent to which the effects of criterion level on associative memory are symmetric remains an open question.

Evidence for Asymmetric Effects of Criterion Level on Associative Memory

Accordingly, in a recent series of experiments, Vaughn and Rawson (resubmitted) minimized concerns over potential differences in the number of pre-existing associates and pre-experimental item accessibility by using pairs of concrete English nouns (i.e., the same stimulus category) and by counterbalancing the assignment of the two words in each pair to serve as the cue versus the target (e.g., either “queen-journal” or “journal-queen”). Regarding item accessibility, we also estimated differences in overall item accessibility by examining the effects of criterion level on separate measures of target memory and cue memory in addition to forward and backward cued recall. During practice, participants learned English-English word pairs to criterion via forward cued recall practice (A → ?). Criterion level (i.e., the number of correct recalls) ranged from one to five. The practice phase for a given participant ended when all the
items had reached their pre-assigned criterion level. The final test occurred one week later. During the final test, we administered forward and backward cued recall measures, as well as measures of cue memory and target memory (described further below).

To evaluate the extent to which criterion level effects on associative memory are symmetric, we evaluated patterns of forward and backward cued recall performance while also considering levels of target memory and cue memory. To revisit, cued recall measures are not pure measures of associative memory, and thus performance on a cued recall test reflects associative memory and also reflects memory for the information being directly retrieved. Accordingly, equivalent gains on the forward versus backward cued recall tests would not necessarily indicate symmetry unless the gains for target memory and cue memory were also equivalent. Likewise, greater performance on the forward versus backward cued recall tests would not necessarily indicate asymmetry if greater gains for target memory versus cue memory were also observed. Thus, evidence that the effects of criterion level on associative memory are symmetric would involve either (1) equivalent gains in forward and backward cued recall and equivalent gains in target memory and cue memory, or (2) greater gains on the forward versus backward cued recall test and greater gains for target memory versus for cue memory. Any other combination of outcomes would indicate that the effects of criterion learning on associative memory are asymmetric given that the differences in forward versus backward cued recall would not be attributable to differences in target versus cue memory. If differences in forward versus backward cued recall cannot be attributed to differences in target or cue memory, then these differences must reflect differences in associative memory.

Results from Vaughn and Rawson (resubmitted) demonstrated that the effects of criterion level on associative memory are asymmetric. In Experiment 1, gains favored forward versus
backward cued recall (see Panel A of Figure 1). However, gains in free recall also favored target versus cue memory (Panel B of Figure 1), suggesting that the effects of criterion learning on associative memory are symmetric. However, gains did not favor target versus cue memory on recognition measures, providing preliminary evidence against associative symmetry (see Panel C of Figure 1). We hypothesized that the discrepancy between free recall and recognition may be due to the possibility that free recall is influenced by associative memory (e.g., an individual may intentionally or unintentionally retrieve some of the cue words and then use the associative links to access the associated targets or vice versa). Furthermore, if the associative link is weaker in the backward direction, participants would be less likely to access cue words from retrieved targets than targets from retrieved cues.

To explore this possibility, we administered only one free recall test in which participants were instructed to recall as many words as possible (i.e., both cues and targets) in Experiment 2. We replicated Experiment 1 by showing greater gains in forward versus backward cued recall (see Panel D of Figure 1). However, gains for targets and cues on the free recall measure did not differ statistically (see Panel E of Figure 1). For recognition, we replicated Experiment 1 by showing numerically greater gains for cues versus targets (the opposite trend necessary for a pattern of associative symmetry) (see Panel F of Figure 1).
FIGURE 1. Mean final test performance as a function of criterion level and final test format in Experiment 1 (left column) and Experiment 2 (right column) from Vaughn and Rawson (resubmitted). Performance is reported in percentages, and error bars report standard error of the mean.
We conducted several conditional analyses to investigate the possible influence of associative memory on free recall. First, when participants freely recalled both words within a pair (e.g., recalling both ‘queen’ and ‘journal’ for the word pair *queen-journal*), secondary analyses revealed that a significantly greater number of word pairs were recalled adjacently (e.g., ‘queen’ followed immediately by ‘journal’) versus non-adjacently (e.g., other items were recalled between ‘queen’ and ‘journal’). These results suggest that associative memory contributed to free recall. Second, the influence of associative memory on free recall was asymmetric, as cues were followed immediately by targets more often than targets were followed immediately by cues (36 vs. 22%, respectively). Taken together, these outcomes indicated that the free recall measure of cue and target memory was influenced by associative memory. Accordingly, we appealed to the recognition measures as better indicators of cue and target memory. In Experiments 1 and 2, we showed numerically greater gains for cue memory versus target memory on the recognition measures, providing further evidence for associative asymmetry (as numerically greater gains in cue versus target memory cannot explain the pattern observed in cued recall) (see Panel C and F of Figure 1). In Experiment 3, we replicated this pattern of performance on cue and target recognition tests that were not preceded by a recall test (to rule out concerns that the recall tests had spillover effects on recognition performance).

The conclusion that the effects of criterion level on associative memory are asymmetric rests in part upon the consistent finding that the effects of criterion level do not differ for cue and target recognition (i.e., no significant interaction between criterion
level and recognition test format). However, a potential concern is that the absence of an interaction between criterion level and recognition performance may reflect scale dependency (e.g., Loftus, 1978). In brief, recognition performance presumably reflects the state of an underlying memory representation. Although test performance and the underlying memory representation are assumed to be monotonically related, the specific function relating these two is not known (e.g., linear, curvilinear, sigmoidal). If the function is not linear, then more caution is needed when interpreting interactions or the lack thereof when comparing test performance in two conditions that are at different parts of the scale (see Panels C and F of Figure 1; note that cue and target recognition are consistently at different parts of the scale). For example, 15% improvements in the lower versus upper parts of a performance scale (e.g., from 5 to 20% versus from 80 to 95%) may not necessarily reflect the same degree of improvement in the underlying memory representations. In contrast, comparisons that involve overlap of performance at some point on the scale support more confident conclusions about the underlying memory representations. Therefore, in Experiment 4 of Vaughn and Rawson (resubmitted; reported here as Experiment 1), we minimized concerns over scale dependency by implementing a pre-exposure phase. During pre-exposure, participants received three blocks of study trials for half the cue words. To foreshadow outcomes described in more detail below, the pre-exposure phase improved cue recognition performance enough that cue and target recognition performance overlapped (reducing concerns over scale-dependency by showing the same pattern at similar parts of the scale). Given these results, I provided pre-exposure for all of the cue words in the present experiments.
(Experiments 2 and 3) to minimize scale dependency concerns on the recognition measures.

Overview of Current Research

The primary question of interest in the proposed research is to ascertain why the effects of criterion learning on associative memory are asymmetric. In order to evaluate why criterion learning produces associative asymmetry, the proposed research was designed to explore the extent to which mediator use can provide a viable account for the consistently observed patterns of associative asymmetry.

According to several modern testing effect theories, the benefits of retrieval practice can be explained via mediator use (mediator refers to any word, phrase, or image which helps link two items in memory). For example, according to the elaborative retrieval hypothesis (ERH; e.g., Carpenter, 2009), attempting retrieval from a cue word activates related semantic information that can be encoded along with the retrieved target, which in turn can later serve as additional pathways to retrieve the target information. For instance, testing on the word pair morning – light (morning - ?) may activate associates of the cue word (e.g., breakfast, shower, commute, tired), all of which can be used later in the service of retrieving the target information (e.g., light). Presumably, the information activated during testing may be more strongly associated with the word serving as the cue (e.g., morning) versus the target (e.g., light), suggesting that the effects of criterion level on associative memory will be asymmetric (favoring forward versus backward associative memory).
Similarly, the *mediator effectiveness hypothesis* (MEH; e.g., Pyc & Rawson, 2010) states that retrieval practice benefits memory by promoting the use of more effective mediators (i.e., information linking cues to targets). For example, keyword mediators commonly involve words that share phonetic similarities to a particular cue in a cue-target word pair (e.g., *wing* for the Swahili-English word pair *wingu – cloud*). The mediator *wing* can then be used to link the mediator to the target word (e.g., birds have wings and fly in the clouds; *wingu - wing - cloud*). Importantly, these verbal mediators are likely to function asymmetrically (e.g., *wingu* elicits *wing*, but *cloud* does not elicit *wing* due to lack of phonetic similarity). If the benefits of testing occur due to more effective mediator use and these mediators tend to link cue-target information asymmetrically, the implication is that the effects of criterion level on associative memory will be asymmetric (again favoring forward versus backward associative memory).

To evaluate this hypothesis, Pyc and Rawson (2010) had participants form keyword mediators to link Swahili-English word pairs (e.g., using the word *wing* to remember that the Swahili word *wingu-cloud*). All participants received an initial block of study trials for 48 Swahili-English word pairs (e.g., *wingu - cloud*), followed by three blocks of practice trials for all the word pairs. In the test-restudy group, each practice trial involved a practice cued recall test over the Swahili-English word pairs (e.g., *wingu - ??*) followed by restudy (e.g., *wingu - cloud*). In the restudy-only group, participants did not engage in retrieval practice and instead received three additional study trials for each
word pair (e.g., *wingu - cloud*). All participants generated and reported keyword mediators during initial study and on each subsequent restudy trial.

Pyc and Rawson (2010) outline two key factors that determine mediator effectiveness: mediator retrieval (i.e., ability to successfully recall a previously generated mediator) and mediator decoding (i.e., degree to which the mediator elicits the target information). To examine mediator retrieval and mediator decoding, Pyc and Rawson (2010) assessed final cued recall performance across three different groups. In the *cue-only* group, participants were presented with a cue word (e.g., *wingu - ???*) and were instructed to type in the corresponding target (e.g., *cloud*). In the *cue + mediator* group, participants were presented with a cue word plus their original mediator (e.g., *wingu – wing - ???*) and were instructed to type in the corresponding target (e.g., *cloud*). Finally, in the *cue + mediator recall* group, participants were presented with a cue word (e.g., *wingu - ??? - ???*) and were instructed to recall both the mediator they generated for that item previously (e.g., *wing*), as well as the corresponding target word (e.g., *cloud*). In all three final test groups, final recall performance favored the test-restudy versus restudy-only group. For the *cue-only* group, this represents a typical testing effect. Of greater interest, other outcomes provided evidence that more effective mediator use contributed to this testing effect. Concerning evidence for enhanced mediator retrieval, participants in the cue + mediator recall group recalled more of their mediators in the test-restudy versus restudy-only group (e.g., *wingu – ??? - ???*). Concerning evidence for enhanced mediator decoding, participants in the cue + mediator group correctly recalled more targets when provided with their original mediator in the test-restudy versus restudy-only
group (suggesting that participants are better able to elicit the correct target word when given their mediator at time of test after test-restudy practice versus restudy-only practice). Pyc and Rawson (2010) interpreted these results as support for MEH as a mechanism to explain the benefits of retrieval practice.

The results from Pyc and Rawson (2010) have direct implications for why criterion learning might have asymmetric effects on associative memory. If the benefits of retrieval practice can be explained at least in part via effective mediator use, then a logical starting point is to investigate whether mediators promote associative symmetry or associative asymmetry. To revisit, for a keyword mediator to be effective, the cue word needs to elicit the mediator (mediator retrieval) and the mediator needs to elicit the target word (mediator decoding). Mediator retrieval is enhanced if the words are phonologically similar (e.g., *wingu* would likely cause one to think of *wing*), whereas mediator decoding is presumably enhanced if the mediator can be related to the target word (e.g., *wing* is related to *cloud* because birds have wings which they use to fly through the clouds). As mentioned previously, keyword mediators would presumably more greatly benefit forward versus backward cued recall due to enhanced mediator retrieval given the greater similarity of the keyword to the cue word versus the target (e.g., *wingu* elicits *wing*, but *cloud* does not elicit *wing* due to lack of phonetic similarity). Relatedly, keyword mediators may more greatly benefit forward versus backward cued recall given that the mediators must first be correctly recalled before they can be correctly decoded to elicit the target word. If participants cannot activate their mediator, then the results from Pyc and Rawson (2010) suggest a decrease in overall final recall.
performance. If the probability of mediator activation is higher on the forward versus backward cued recall tests, mediator use may provide at least a partially viable explanation for the observed pattern of associative asymmetry.

Thus far, I have discussed how mediators may at least partially account for the observed pattern of associative asymmetry. However, in order for mediator use to be a viable account, I need to establish that participants are indeed using mediators spontaneously during practice. In contrast to Pyc and Rawson (2010), participants in the prior experiments demonstrating associative asymmetry were not instructed to form mediators during practice. As such, there are several unknowns which need to be explored for a mediator-based explanation to be a viable account of these findings of associative asymmetry. The first issue concerns how often participants report using mediators without any explicit instruction to do so. The second issue concerns the possibility that participants could be forming other kinds of mediators during practice (e.g., visual mediators). An example of a visual mediator might be an interactive image (e.g., visualizing a dog eating with a spoon for the word pair dog-spoon). Presumably, visual mediators would not be subject to the same asymmetric limitations as would keyword mediators due to the fact that an image contains both the cue word (e.g., dog) and the target word (e.g., spoon). If the image contains both cue and target elements, then the visual mediator would presumably be activated on both a forward and backward cued recall test. If both the cue word and the target word can elicit the visual mediator, then items associated via a visual mediator may produce patterns more consistent with
associative symmetry (e.g., similar performance on forward versus backward cued recall measures).

The notion that visual mediators may produce patterns more consistent with associative symmetry is motivated by findings within the literature on associative memory and mediator use. For instance, Wollen and Lawry (1971) found similar performance on measures of forward versus backward cued recall when participants used interactive imagery to relate unrelated English-English word pairs (e.g., visualizing a tree wrapped in carpet for the word pair carpet – tree; for similar outcomes, see Mondanni & Battig, 1973). Although these studies suggest that visual mediators may attenuate the pattern of associative asymmetry, there are several methodological differences which make it difficult to draw firm conclusions regarding how these studies may inform the present research. For example, neither the Wollen and Lawry (1971) study nor the Mondanni and Battig (1973) study manipulated criterion level; to my knowledge, no prior studies investigating mediator use have manipulated criterion level. Furthermore (and perhaps most important for present purposes), prior studies investigating mediator use have assessed associative symmetry primarily in the context of comparing forward versus backward cued recall, which is problematic because cued recall measures reflect associative memory as well as the memory for the information being directly retrieved (as outlined above).

Therefore, the purpose of the proposed research is to answer both empirical and theoretical questions regarding why criterion level produces asymmetric gains in associative memory. First, Experiment 1 was conducted to demonstrate the pattern of
associative asymmetry across a series of recall and recognition measures. Second, Experiment 2 was conducted to investigate whether participants spontaneously generate mediators during practice, which is a necessary condition for a mediator-based explanation to potentially account for these patterns of associative asymmetry. Finally, I manipulated mediator use in Experiment 3 to provide a more explicit test of how mediator use influences the pattern of associative asymmetry via criterion learning.
EXPERIMENT 1

The purpose of Experiment 1 was to establish that criterion learning produces asymmetric effects on associative memory. Additionally, Experiment 1 was also designed to minimize potential concerns about scale-dependency by including conditions in which cue and target recognition were at similar levels on the performance scale (as discussed above).

Method

Participants and Design

One hundred fifteen undergraduates participated for course credit. Criterion level during practice (1, 3, or 5 correct recalls) was a within-participant manipulation. Item pre-exposure was a within-participant manipulation. Final test format was a between-participant manipulation, and participants were randomly assigned to complete either a forward cued recall test, a backward cued recall test, a cue recognition test, or a target recognition test.

Materials

Materials included 48 unrelated English-English word pairs (e.g., queen-journal), divided into six lists of eight pairs. The cue-target pairs had zero forward and zero backward associative strength (Nelson, McEvoy, & Schreiber, 2004). Which word
served as the cue versus the target was approximately counterbalanced across participants. Assignment of list to pre-exposure condition and criterion level was approximately counterbalanced across participants.

Procedure

First, in the pre-exposure phase at the beginning of the experiment, cue words from half of the word pairs (24 cue words) were presented for an initial study trial one at a time for 4 seconds each. After the first block of pre-exposure trials, the procedure was repeated for two additional blocks of pre-exposure trials. Presentation order was randomized during each block of pre-exposure trials. Next, all 48 word pairs were presented for initial study in random order one at a time via computer for 10 seconds each. The word on the left was always presented in blue font under the heading “cue word,” whereas the word on the right was always presented in red font under the heading “target word.” After all items were studied once, the practice phase began. During practice trials, participants were shown a cue word (always presented in blue font on the left) and had up to 8 seconds to type in the corresponding target. This test format represents forward cued recall (A - ?????) and was the only cued recall test format used during practice. After incorrect responses, the cue and target were restudied for 4 seconds. The item was then placed at the end of the list for another practice trial later. If a response was correct but the item had not yet reached its assigned criterion, it was placed at the end of the list for another practice trial later. Items were dropped from
practice once they reached their pre-assigned criterion. Session 1 ended when all items reached criterion or after 75 minutes had elapsed. Data for one participant who did not reach at least 75% criterion for all items during the allotted time in Session 1 were excluded from analyses.

In Session 2 seven days later, participants completed one of four final tests depending on group assignment. For all final test groups, items were presented one at a time and presentation order was randomized anew for each participant. For the group of participants who completed forward cued recall, cues were presented in blue font on the left, and participants were instructed to type in the corresponding target (e.g., queen - ???). For the group of participants who completed backward cued recall, targets were presented in red font on the right and participants were instructed to type in the corresponding cue (e.g., ??? - journal). For the cue recognition and target recognition groups, the 48 previously learned English words (either the cue words or target words, respectively) were randomly mixed with 48 novel English words. Participants were instructed to indicate whether the presented word on the screen represented a previously learned word. In all final test groups, participants had an unlimited amount of time to respond.
Results and Discussion

To revisit, evidence that the effects of criterion level on associative memory are symmetric would involve either (1) equivalent gains in forward and backward cued recall and equivalent gains in target memory and cue memory, or (2) greater gains on the forward versus backward cued recall test and greater gains for target memory versus for cue memory. First, consider forward versus backward cued recall performance (see Figure 2). A 2 (pre-exposure or no-pre-exposure) X 2 (type of cued recall test) X 3 (criterion level) mixed factor ANOVA revealed a main effect of criterion level \( [F(2,114) = 85.16, MSE = 1.32, p < .001, \eta^2_p = .60] \) and a significant 2-way interaction between type of cued recall test (forward vs. backward) and criterion level \( [F(2,114) = 8.89, MSE = 1.32, p < .001, \eta^2_p = .14] \). Thus, there were greater gains in forward versus backward cued recall (32 versus 16%). There was no main effect for pre-exposure condition \( [F(1,57) = 1.40, MSE = 1.31, p = .242, \eta^2_p = .02] \) nor any interactions involving pre-exposure condition \( [Fs < 1.41] \).
FIGURE 2. Mean final cued-recall performance as a function of criterion level, type of cued recall test, and pre-exposure condition in Experiment 1. Performance is reported in percentages, and error bars report standard error of the mean.

Recognition outcomes are reported in Figure 3. A 2 (pre-exposure or no-pre-exposure) X 2 (type of recognition test) X 3 (criterion level) mixed factor ANOVA revealed a main effect of criterion level \(F(2,106) = 39.41, \text{MSE} = .02, p < .001, \eta^2_p = .43\) and pre-exposure condition \(F(1,53) = 4.57, \text{MSE} = .03, p = .037, \eta^2_p = .08\).

Additionally, the interaction between pre-exposure condition and type of recognition test was significant \(F(1,53) = 5.64, \text{MSE} = .03, p = .021, \eta^2_p = .10\), as well as the interaction between pre-exposure condition and criterion level \(F(2,106) = 5.37, \text{MSE} = .02, p = .006, \eta^2_p = .09\). There was also a significant 3-way interaction, \(F(2,106) = 4.76, \text{MSE} = .02, p = .010, \eta^2_p = .08\) (all other Fs < 3.62). To explore the 3-way interaction, I
conducted separate 2 (type of recognition test) x 3 (criterion level) ANOVAs for the pre-exposure and no pre-exposure conditions. The interaction between type of recognition test (cue vs. target) and criterion level was significant in the no-pre-exposure condition \[F(2,106) = 5.10, MSE = .02, p = .008, \eta^2_p = .09\], but not in the pre-exposure condition \[F < 1\]. These results confirm greater gains for cue memory versus target memory on the recognition measures in the no-pre-exposure condition (26 versus 11%) and no significant difference in gains in the pre-exposure condition (11 versus 15%), neither of which is consistent with a symmetry account.

FIGURE 3. Mean final recognition performance as a function of criterion level, type of recognition test, and pre-exposure condition in Experiment 1. Performance is reported in percentages, and error bars report standard error of the mean.
In sum, Experiment 1 showed that gains were greater in forward versus backward cued recall. Importantly, gains favored forward versus backward cued recall regardless of whether cue memory and target memory gains were equivalent (in the pre-exposure condition) or whether gains were greater for cue versus target memory (in the no pre-exposure condition). Therefore, the differential gain in forward versus backward cued recall performance cannot be attributed to differential gain in target versus cue memory, providing evidence that the effects of criterion level on associative memory are asymmetric.
Experiment 2

Experiment 1 showed that criterion level produces asymmetric gains on associative memory. The purpose of Experiment 2 is to explore to what extent mediator use may be able to account for the observed pattern of associative asymmetry. To this end, I investigated the frequency and type of mediators participants report using spontaneously (i.e., without explicit instruction to generate mediators during practice). Within the literature on mediator use, researchers have assessed mediator use concurrently or retrospectively (e.g., Dunlosky & Hertzog, 1998, 2001; Richardson, 1998). Concurrent reports of mediator use involve having participants report their mediator immediately after studying a given item. In contrast, retrospective reports of mediator use involve having participants report their mediator after a delay. Proponents of concurrent reports claim that reporting mediators immediately leads to higher accuracy in reported mediator use given a minimal degree of forgetting (as might occur in retrospective reporting). Proponents of retrospective reports claim that participants are not able to change their encoding strategies with reports made after a delay (i.e., minimal reactive effects, which might occur in concurrent reporting).

Recent research by Dunlosky and Hertzog (2001) investigated mediator use both concurrently and retrospectively on an item-by-item basis across three groups: concurrent, retrospective informed, and retrospective uninformed. For both the concurrent and retrospective informed groups, participants received instructions about the
various types of mediators commonly used (e.g., sentence generation) along with a one sentence description (e.g., “when you use sentence generation, you try to link the two words together by completing a sentence that includes both words”) as well as two examples of each mediator strategy. In the retrospective uninformed group, no information on mediators was given to the participants until after they had studied and been tested on all the items (i.e., to avoid any possible reactive effects). For all groups, participants indicated which mediator they used for that particular item from six possible choices: interactive imagery, rote repetition, sentence generation, other strategy, no strategy, or an option to indicate that they had tried to use a strategy but ran out of time. For the concurrent group, participants indicated which mediator they had used both immediately after studying that particular item and after all items had been studied and tested (i.e., during initial study and at the end of the experiment). For the retrospective groups, participants indicated which mediator they had used during study only after all items had been studied and tested (i.e., at the end of the experiment).

For each participant, reported mediator strategies were analyzed on an item-by-item basis. In the concurrent group, mediators reported retrospectively tended to match those reported concurrently (i.e., participants remained consistent with the frequency and type of mediators reported across the two different strategy reports). Furthermore, the retrospective reports made by the informed retrospective group were similar to those made by the uninformed retrospective group (72 vs. 74% reported mediator use across trials, respectively), suggesting minimally reactive effects. Finally, proportion correct on
the cued recall test administered at the end of the experiment did not differ between the three groups. Overall, these results suggest that participants report similar mediator information and exhibit similar final test performance regardless of whether or not they receive instructions on mediator use prior to the start of the experiment and also regardless of whether they report mediators concurrently or retrospectively. Given that I am most interested in participants’ spontaneous mediator use, I assessed mediator strategy use retrospectively on an item-by-item basis (without giving prior instructions on mediators, as in the uninformed retrospective group described in Dunlosky & Hertzog, 2001).

Method

Participants and Design

Eighty-six undergraduates participated for course credit. Criterion level during practice (1, 3, or 5 correct recalls) was a within-participant manipulation. In contrast to Experiment 1, all items received pre-exposure in Experiment 2. All participants completed retrospective strategy reports at the end of Session 1. Final test format was a between-participant manipulation, and participants were randomly assigned to complete either a forward cued recall test, a backward cued recall test, a cue recognition test, or a target recognition test.

Sample size was determined using power analyses conducted with the statistical program G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). Power analyses were
conducted based on the analysis plan to conduct two separate 2 (type of final test) x 3 (criterion level) repeated measures mixed factor ANOVAs, one involving the cued recall groups and the other involving the recognition groups. Based on power at the .90 level to detect a within-between interaction of moderate effect size (.25), G*Power determined the appropriate sample size to be 44 participants for each 2x3 mixed factor ANOVA (resulting in a final sample size of 88 participants).

Materials and Procedure

Materials and procedure were identical to Experiment 1, with the following exceptions. First, all items received pre-exposure in order to minimize interpretive difficulty arising from potentially scale-dependent interactions. Second, I assessed spontaneous mediator use retrospectively after all items had been learned to criterion in Session 1 (without informing participants that mediator use would be assessed at the end of the session, as in the uninformed retrospective group in Dunlosky & Hertzog, 2001). As in Dunlosky and Hertzog (2001), I presented all participants with instructions on mediator use prior to the mediator assessment reports at the end of the session. Each participant read instructions on the different kinds of mediators one might use to associate two items in memory (e.g., sentence generation), along with a brief description (e.g., “when you use sentence generation, you try to link two words together using a sentence”) and two examples (e.g., “For the word pair “bug-plant,” you might say, “the bug crawled slowly up the plant”). During the mediator assessment reports, participants
were presented with each word pair and given seven choices for the potential mediator
types they may have used (along with reminder instructions, listed in parentheses):
imagery (“when you form a mental picture in your mind that combines the two words”),
repetition (“when you repeat information over and over”), sentence (“when you link two
words using a sentence”), keyword (“when you link two words use a related word”),
other (“you mostly used something other than one of these four”), none (“you didn’t ever
try to use a strategy”), or ran out of time (“you tried to use a strategy to learn a pair but
didn’t have enough time on the practice trials to fully form one”). Participants were
instructed to indicate which mediator they had used most often during practice to learn
that word pair. The order of items was randomized anew for each participant, and
participants had an unlimited amount of time to indicate the mediator they used most for
each word pair. As a potential safeguard to prevent participants from repeatedly clicking
one button over and over, the order of the seven choices was randomized anew on an
item-by-item basis. Data were excluded for one participant who did not reach at least
75% criterion.

Results and Discussion

First, consider cued recall outcomes (see Figure 4). A 2 (type of cued recall test)
X 3 (criterion level) mixed factor ANOVA replicated Experiment 1 with a main effect of
criterion level \( F(2, 80) = 19.80, \text{MSE} = 141.34, p < .001, \eta^2_p = .33 \) and a significant 2-
way interaction between type of cued recall test (forward vs. backward) and criterion
level (i.e., greater gains in forward versus backward cued recall, 21 vs. 9%, respectively) 
$[F(2,80) = 3.40, MSE = 141.34, p = .038, \eta_p^2 = .08]$. There was no main effect of final
test format [$F < 1$]. These results are consistent with the cued recall outcomes reported in
Experiment 1 and are also consistent with the anticipated pattern of associative
asymmetry.

**FIGURE 4.** Mean final cued-recall performance as a function of criterion level and type
of cued recall test in Experiment 2. Performance is reported in percentages, and error
bars report standard error of the mean.

Next, consider recognition outcomes (see Figure 5). A 2 (type of recognition test)
X 3 (criterion level) mixed factor ANOVA replicated Experiment 1 by showing a main
effect of criterion level $[F(2,82) = 22.95, MSE = 99.92, p < .001, \eta_p^2 = .36]$. In contrast to
the pre-exposure group in Experiment 1, there was also a main effect of recognition test format \[ F(1,41) = 7.64, MSE = 391.29, p = .009, \eta^2_p = .16 \]. The main effect of recognition test format suggests that the pre-exposure manipulation did not produce equivalent cue and target recognition as in Experiment 1. It is unclear why differences in recognition performance emerged in Experiment 2 but not in Experiment 1 given that the same pre-exposure methodology was used in both experiments. Finally, although the numerical trend showed somewhat greater gain for target recognition than for cue recognition (16 vs. 10%, respectively), the interaction between type of recognition test (cue vs. target) X criterion level was not significant \[ F(2,82) = 1.93, MSE = 99.92, p = .151, \eta^2_p = .05 \]. Together, these cued recall and recognition results suggest that the advantage of forward versus backward cued recall is unlikely to be completely attributable to differences in target and cue accessibility, suggesting that associative asymmetry is at least partly responsible for the greater gains in forward versus backward cued recall.
FIGURE 5. Mean final recognition performance as a function of criterion level and type of recognition test in Experiment 2. Performance is reported in percentages, and error bars report standard error of the mean.

To investigate the extent to which mediators may explain associative asymmetry, the primary goal of Experiment 2 was to establish that participants use mediators spontaneously during practice. To this end, I calculated the percentage of items for which participants reported using a mediator (sentence, keyword, or imagery) during initial learning. I did not include trials in which participants indicated using “repetition”; although this represents a learning strategy, it does not qualify as a mediator (i.e., saying things over and over is not the same as linking two items together through some sort of visual or verbal mnemonic). Based upon the proportions of mediator use reported by
Dunlosky and Hertzog (2001) for the uninformed retrospective group (74% mediator use across trials), I predicted that participants would report using a mediator approximately 70-75% of the time. In line with this prediction, participants reported using a mediator for the majority of items ($M = 73\%$, $SEM = 2.06$), and significantly greater than would be expected if participants were randomly selecting response options [3/7 options = 43%, $t(84) = 14.58, p < .001, d = 3.18$]. Therefore, results indicate that participants do use mediators spontaneously during practice, a necessary condition for a mediator-based explanation of asymmetry to be viable.

Finally, I conducted conditional analyses to determine if the same asymmetric pattern emerges with different kinds of mediators (with an emphasis on verbal vs. visual mediators). For verbal mediators, I expected to replicate the same asymmetric pattern observed in Experiment 1 (with greater gains in forward versus backward cued recall; but equivalent gains in cue versus target recognition). For visual mediators, I expected a more symmetric pattern to emerge (with equivalent gains in forward versus backward cued recall; as well as equivalent gains in cue versus target recognition).

Before discussing recall and recognition performance as a function of initial mediator use, it is important to bear in mind that these conditional analyses are subject to item selection effects. In Experiment 2, I did not explicitly manipulate mediator use during practice, and participants therefore self-selected which mediator strategy they used for each item. Furthermore, participants reported using a mix of strategies to learn the items during practice (see Table 1) and therefore only a subset of items can contribute to
each specific mediator strategy (i.e., verbal or visual). Also, this subset of items must then be divided further into their appropriate criterion levels. For both verbal and visual mediators, separate 2 (type of final test) x 3 (criterion level) mixed factor ANOVAs for cued recall groups and recognition groups revealed no significant differences in the number of items contributing at each criterion level, in the number of items contributing between final test formats, and no significant interaction between criterion level and final test format (all $F$s < 2.64). Thus, participants used verbal and visual mediators equally often regardless of the initial criterion level of the item and regardless of their final test format (which is perhaps not surprising given that they were not aware of the type of final test they would be taking at the end of Session 1, when the mediator assessment phase occurred).
<table>
<thead>
<tr>
<th>Strategy Type</th>
<th>Sentence</th>
<th>Keyword</th>
<th>Imagery</th>
<th>Repetition</th>
<th>Other</th>
<th>None</th>
<th>Ran out of time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 2</td>
<td>28.6 (2.4)</td>
<td>8.8 (1.1)</td>
<td>31.8 (2.2)</td>
<td>21.4 (2.0)</td>
<td>3.8 (0.7)</td>
<td>3.6 (0.8)</td>
<td>1.9 (0.4)</td>
</tr>
<tr>
<td>Experiment 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mediator Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td>42.1 (2.7)</td>
<td>6.0 (0.8)</td>
<td>32.1 (2.5)</td>
<td>12.3 (1.5)</td>
<td>3.2 (0.9)</td>
<td>2.7 (0.6)</td>
<td>1.6 (0.4)</td>
</tr>
<tr>
<td>Spontaneous</td>
<td>28.4 (2.7)</td>
<td>9.9 (1.1)</td>
<td>30.4 (2.5)</td>
<td>20.8 (2.2)</td>
<td>4.7 (0.8)</td>
<td>4.1 (0.9)</td>
<td>1.7 (0.4)</td>
</tr>
</tbody>
</table>

TABLE 1. Mean percentage of strategy use reported during the retrospective strategy reports in Experiments 2 and 3.
With the caveats noted above in mind, a 2 (type of test) X 3 (type of mediator used) X 3 (criterion level) mixed factor ANOVA revealed only a significant main effect of criterion level level \( [F(2, 54) = 7.79, \text{MSE} = 720.61, p = .001, \eta_p^2 = .22] \) (all other Fs < 2.62), suggesting that final test performance may not depend on the type of mediator used (given the lack of a significant interaction between final test performance and type of strategy used during practice). However, caution is needed when interpreting these results given that these analyses excluded participants who did not have at least one contributing performance value for each strategy type (i.e., visual mediator, verbal mediator, and “other” strategy) across each criterion level (thus reducing the number of participants contributing to the ANOVA).

To further explore how strategy use influences final test performance (and to increase the number of participants contributing to each ANOVA), I conducted separate 2 (type of test) X 3 (criterion level) mixed factor ANOVAs for each type of strategy (i.e., verbal mediator, visual mediator, and “other” strategy). For cued recall performance with verbal mediators (i.e., sentence and keyword mediators; see the left panel of Figure 6), a 2 (type of test) X 3 (criterion level) mixed factor ANOVA revealed a main effect of criterion level \( [F(2,76) = 5.47, \text{MSE} = 480.68, p = .006, \eta_p^2 = .13] \) but no significant interaction between criterion level and final test format \( [F(2,76) = 1.42, \text{MSE} = 480.68, p = .248, \eta_p^2 = .04] \). Also, there was no significant main effect of final test format \( [F < 1] \). The lack of a significant interaction between criterion level and final test format (as well as the lack of a performance difference between forward and backward cued recall) is not
consistent with the asymmetric pattern I predicted for items encoded with verbal mediators. In contrast, these results provide suggestive evidence that verbal mediators do not necessarily promote associative asymmetry given the similar gains across criterion level for both forward and backward cued recall.
FIGURE 6. Mean final cued-recall performance for items in which participants reported using a verbal mediator during practice (left panel), a visual mediator during practice (middle panel), or some other kind of strategy (including no strategy or repetition) during practice (right panel), represented as a function of criterion level and type of cued recall test in Experiment 2. Performance is reported in percentages, and error bars report standard error of the mean.
Next, consider cued recall outcomes for items that participants reported using visual mediators to learn during practice (i.e., a mediator involving imagery; see the middle panel of Figure 6). For cued recall performance with visual mediators, a 2 (type of test) X 3 (criterion level) mixed factor ANOVA revealed a main effect of criterion level \([F(2,68) = 10.80, \text{MSE} = 410.71, p < .001, \eta^2_p = .24]\) but no significant interaction between criterion level and final test format \([F(2,68) = 1.19, \text{MSE} = 410.71, p = .310, \eta^2_p = .03]\). Again, there was no significant main effect of final test format \([F < 1]\). The lack of a significant interaction between criterion level and final test format (as well as the lack of a performance difference between forward and backward cued recall) is consistent with the symmetric pattern I predicted for items encoded with visual mediators.

For items that participants reported using a non-verbal or non-visual strategy to learn during practice (i.e., a repetition strategy, no strategy, or some other kind of strategy; see the right panel of Figure 6), I made no a priori predictions for this subset of items, but for purposes of completeness, a 2 (type of test) X 3 (criterion level) mixed factor ANOVA revealed no main effect of criterion level, no main effect of final test format, and no significant interaction between criterion level and final test format [all Fs < 2.36]. The lack of a significant interaction between criterion level and final test format (as well as the lack of a performance difference between forward and backward cued recall) is consistent with the symmetric pattern that obtained for items encoded with verbal or visual mediators.

For cued recall, these outcomes provide suggestive evidence that verbal mediator use may not promote associative asymmetry (i.e., participants experience similar forward
and backward recall gains across criterion level regardless of whether they used a verbal or visual mediator during practice). Rather, patterns consistent with an associative symmetry account emerged for cued recall for both verbal and visual mediators.

Next, consider recognition outcomes for items that participants reported using verbal mediators to learn during practice (i.e., sentence and keyword mediators; see the left panel of Figure 7). For recognition performance with verbal mediators, a 2 (type of test) X 3 (criterion level) mixed factor ANOVA revealed a main effect of criterion level \( [F(2,74) = 4.39, \text{MSE} = 398.73, p = .016, \eta^2_p = .11] \) but no significant interaction between criterion level and final test format, as well as no main effect of final test format [both \( F \)'s < 1]. Again, these results are consistent with an associative symmetry account. When a verbal mediator was used to learn items during practice, both cue and target recognition gains were similar across criterion level. Furthermore, cue and target recognition performance were not significantly different. These recognition outcomes are consistent with the cued recall outcomes reported previously, suggesting that verbal mediators may not be able to explain the previously observed pattern of associative asymmetry.
FIGURE 7. Mean final recognition performance for items in which participants reported using a verbal mediator during practice (left panel), a visual mediator during practice (middle panel), or some other kind of strategy (including no strategy or repetition) during practice (right panel), represented as a function of criterion level and type of recognition test in Experiment 2. Performance is reported in percentages, and error bars report standard error of the mean.
Next, consider recognition outcomes for items that participants reported using visual mediators to learn during practice (i.e., a mediator involving imagery; see the middle panel of Figure 7). For recognition performance with visual mediators, a 2 (type of test) x 3 (criterion level) mixed factor ANOVA revealed a main effect of criterion level \( F(2,70) = 5.71, \text{MSE} = 260.79, p = .005, \eta_p^2 = .14 \) but no significant interaction between criterion level and final test format, as well as no main effect of final test format [both \( F \text{s} < 1 \)]. Again, there was no significant main effect of final test format \( [F < 1] \).

For items that participants reported using a non-verbal or non-visual strategy to learn during practice (i.e., a repetition strategy, no strategy, or some other kind of strategy; see the right panel of Figure 7), a 2 (type of test) x 3 (criterion level) mixed factor ANOVA revealed a main effect of criterion level \( F(2,70) = 11.52, \text{MSE} = 231.61, p < .001, \eta_p^2 = .25 \), but no main effect of final test format and no significant interaction between criterion level and final test format \( [F \text{s} < 1.20] \). The lack of a significant interaction between criterion level and final test format (as well as the lack of a performance difference between cue and target recognition) is consistent with the symmetric pattern that obtained for items encoded with verbal or visual mediators.

**Summary**

Experiment 2 revealed several important findings. First, cued recall performance gains across criterion level were greater in the forward versus backward direction. For recognition, performance gains numerically favored target recognition versus cue recognition; however, the interaction between criterion level and recognition test format was not significant. Therefore, the greater gains in forward versus backward cued recall
were unlikely to be completely attributable to differences in cue or target memory gains, suggesting that there were differences in forward versus backward associative memory gains.

Second, results from Experiment 2 indicate that participants used mediators spontaneously (i.e., with no explicit instructions to do so) during criterion learning, which is a necessary condition for a mediator-based account to be a viable explanation for asymmetric effects of criterion level on associative memory. However, conditional analyses unexpectedly provided suggestive evidence that mediator use promotes associative symmetry. For both verbal and visual mediators, no interactions between criterion level and final test format were observed for cued recall measures or for recognition measures. Additionally, for both verbal and visual mediators, there were no differences between forward versus backward cued recall and no differences between cue versus target recognition. Although exploratory in nature, these patterns suggest that verbal mediators may not offer a viable explanation for the previously observed patterns of associative asymmetry.
Results from Experiment 2 established that participants use mediators spontaneously during criterion learning, which is a necessary condition for mediator use to provide a plausible explanation for why criterion level produces asymmetric gains in associative memory. In addition, Experiment 2 provided suggestive evidence that mediator use (verbal and visual) may not be able to offer a viable explanation of the observed patterns of associative asymmetry. However, as mentioned previously, these results cannot be taken as strong evidence because I did not manipulate mediator use during initial learning (rather, I assessed it after all items had been learned to criterion in Session 1). To provide a clearer examination of how mediator use influences associative symmetry, I directly manipulated mediator use during initial learning in Experiment 3. If mediator use is at least partially responsible for these patterns of associative asymmetry, then I should be able to control the degree of associative asymmetry observed through direct manipulations of mediator use during initial learning. Specifically, I expected to observe patterns consistent with associative asymmetry when verbal mediators are used (i.e., greater forward versus backward cued recall despite similar levels of cue and target recognition performance) and patterns consistent with associative symmetry when visual mediators are used (i.e., similar forward versus backward cued recall as well as similar levels of cue and target recognition performance).
Method

Participants and Design

One hundred fifty-nine undergraduates participated for course credit. Mediator group was a between-participant manipulation (verbal versus spontaneous). As in Experiment 2, all items received pre-exposure and all participants completed retrospective strategy reports at the end of Session 1. Final test format was also a between-participant manipulation (forward cued recall, backward cued recall, or cue and target recognition). In contrast to Experiments 1 and 2, I combined cue and target recognition into one measure in Experiment 3 in order to keep the design more manageable. Additionally, all items were learned to a criterion of 5 (rather than manipulating three different criterion levels within-participant) to increase the number of items available for subsequent mediator analyses in the spontaneous mediator group as in Experiment 2 (i.e., recall and recognition performance conditionalized by mediator type). Based on the results from Dunlosky and Hertzog (2001) and Experiment 2, I expected a mediator to be generated on approximately 70-75% of the trials (which would comprise approximately 28 items; see below in the materials section for a change to the total number of items used). In Experiment 2, participants reported using a visual mediator on approximately one-third of all trials ($M = 31.8\%, \ SEM = 1.06$). If participants report using a visual mediator to the same extent in Experiment 3, then the number of items remains reasonable with only one criterion level (approximately 12 word pairs, versus further subdividing this subset across three criterion levels as in Experiment 2).
Sample size was determined based upon power analyses conducted in the statistical program G*Power (see Faul, Erdfelder, Lang, & Buchner, 2007). Power analyses were based on the assumption that I would conduct two separate 2 (verbal mediator or spontaneous mediator group) x 2 (type of final test) ANOVAs. The ANOVAs are powered at the .80 level to detect an interaction of moderate effect size (.25). For the 2 (verbal mediator or spontaneous mediator group) x 2 (forward or backward cued recall) ANOVA, G*Power determined the appropriate sample size to be 32 participants for each group (resulting in 128 participants required for this ANOVA). For the 2 (verbal mediator or spontaneous mediator group) x 2 (cue or target recognition) mixed factor ANOVA, G*Power determined the appropriate sample size to be seventeen participants for each group. However, based upon recent recommendations for minimum group sizes of 20 participants (Simmons, Nelson, & Simonsohn, 2011), I increased the number of participants in each group to 20 (resulting in 40 participants for this ANOVA; see Simmons et al., 2011 for additional information on sample size and statistical power).

Materials

Materials included 36 unrelated English-English word pairs (e.g., queen-journal) extracted from the same item list as used in Experiment 1 and 2. Given that all items were learned to Criterion 5 in Experiment 3, I used a smaller number of word pairs in order to keep the overall length of the experiment and mean number of trials in Session 1 comparable to Experiments 1 and 2. To create the shortened list of 36 word pairs, I conducted item analyses in Experiment 2 and removed the 12 word pairs which elicited the highest spontaneous verbal mediator use, to increase the number of items for which
participants report using other kinds of mediators (e.g., visual) in the spontaneous group. Increasing the number of items for which participants in the spontaneous group generate non-verbal mediators was intended to reduce the overlap between mediators used in the verbal versus spontaneous mediator groups (i.e., to provide a stronger manipulation of mediator use during practice).

Procedure

First, participants in both mediator groups received three blocks of pre-exposure for all 36 cue words. Then, participants in the verbal mediator group were instructed on verbal mediators (e.g., “these mediators involve linking the cue word to the target word via a word or sentence”). Furthermore, two examples were provided for each mediator strategy type (e.g., for dog – spoon: “the dog eats with a spoon”). Next, participants in the verbal mediator group completed one practice trial for an unrelated word pair (a novel word pair not to be learned in the actual experiment; “baby-phone”). This practice trial was administered to ensure that participants comprehended the instructions regarding verbal mediator use. After participants completed the practice trial (e.g., “the baby talked on the phone”), the experimenter advanced the computer program to the next phase of the experiment. Importantly, participants in the spontaneous mediator group did not receive instructions regarding mediator use before the experiment.

Next, the study phase began. During the study phase, each word pair was presented for initial study in random order one at a time via computer for 10 seconds. After each study trial, participants in the verbal mediator group were given 10 seconds to type in the verbal mediator they used for that specific item (participants in the
spontaneous mediator group only reported mediators after all items were learned to criterion, as in Experiment 2). After the study phase, the criterion learning phase was administered as in Experiments 1 and 2, except that all items were practiced to a criterion of five correct recalls. In order to ensure that participants in the verbal mediator group were following instructions and using verbal mediators to associate the word pairs, to the computer initiated a mediator report trial at random on approximately 10% of the trials. During the mediator report trials, participants in the verbal mediator group were presented with a mediator report screen. On the mediator report screen, both the cue and target word were presented along with a response box (with instructions for participants to type in the mediator they had used for that particular word pair). Participants had 10 seconds to type in their verbal mediator. If they finished before the 10 seconds elapsed, they were able to click a button labeled “done”. After each mediator report trial, the computer recorded their typed response and advanced to the next practice trial.

After all items had reached criterion, all participants completed retrospective strategy reports indicating the strategy they had used on an item-by-item basis. As in Experiment 2, all participants were first given information about mediator strategies (similar to the information the participants in the verbal mediator group received prior to the start of Session 1, except that information was provided on visual mediators, verbal mediators, and rote repetition). For the verbal mediator group, instructions were tailored to encourage participants to accurately and honestly indicate which mediator they used for each specific item during practice (“Earlier, we instructed you to form a verbal mediator for each word pair. However, we understand that you may not have always
used a verbal mediator during practice. For each pair, we’d like you to tell us the strategy you actually used most often during practice to learn that pair.”). Participants then reported the strategies they had used for each item using the same procedure as in Experiment 2. Session 1 ended after participants had reported their strategy for each item or after 90 minutes had elapsed. Data were excluded for four participants who did not reach at least 75% criterion during Session 1.

Session 2 occurred after a 1-week delay and followed a similar format as in Experiment 2. Participants completed either a forward cued recall test, a backward cued recall test, or a recognition test (comprising all 36 cue and target words, intermixed with seventy-two novel lures). On all tests, items were randomized anew for each participant and presented one at a time. Participants had an unlimited amount of time to respond.

Results and Discussion

Before turning to primary outcome measures, I will briefly report outcomes relevant to compliance with instructions. First, I analyzed how often participants in the verbal mediator group generated a verbal mediator during the initial mediator report trials. Importantly, participants in the verbal mediator group generated a verbal mediator on 98.2% (SEM = .82) of the initial 36 mediator report trials. To further ensure that participants in the verbal mediator group were using verbal mediators during practice, the computer program was designed to administer a mediator report trial at random on approximately 10% of trials throughout the learning phase of Session 1 (actual percentage of randomized mediator report trials during the test phase was 12.5% [SEM = .10]). Of interest, participants generated a verbal mediator on 91.5% (SEM = 2.00) of the
randomly occurring mediator report trials that were administered during the learning phase of Session 1. These results confirm that participants in the verbal mediator group complied with the instructions and generated a verbal mediator on the vast majority of both the initial mediator report trials (i.e., the mediator report trials occurring immediately after the word pair was studied) and the mediator report trials occurring during the test phase (i.e., the mediator report trials occurring at random when items were being learned to criterion).

Next, I calculated the percentage of different strategies participants reported using during practice (see Table 1, Experiment 3). As expected, participants reported using more verbal mediators (i.e., sentence and keyword mediators) in the verbal mediator group ($M = 48.2\%$, $SEM = 2.7$) versus the spontaneous group ($M = 38.3\%$, $SEM = 2.7$) [$F(1,149) = 6.85$, $p = .010$, $\eta_p^2 = .04$], providing additional evidence that the mediator group manipulation was successful (as participants used more verbal mediators in the verbal mediator group). Additionally, participants reported using a visual mediator equally often in both the verbal mediator group ($M = 32.1\%$, $SEM = 2.5$) and the spontaneous mediator group ($M = 30.4\%$, $SEM = 2.5$) [$F < 1$].

Concerning primary outcome measures, cued recall results are reported in Figure 8. A 2 (verbal mediator group vs. spontaneous group) X 2 (forward vs. backward cued recall) between-participant ANOVA revealed a main effect of final test format (with greater forward versus backward cued recall performance) [$F(1,113) = 8.01$, $MSE = 565.89$, $p = .006$, $\eta_p^2 = .07$], but no main effect of mediator group [$F < 1$]. Also, the interaction between final test format (forward versus backward cued recall) X mediator
group (verbal mediator group versus spontaneous group) was not significant \( F < 1 \). Thus, participants remembered more words when tested in the forward versus backward cued recall direction regardless of whether or not they were instructed to explicitly generate verbal mediators during practice.

![FIGURE 8](image)

**FIGURE 8.** Mean final cued-recall performance as a function of mediator group and type of cued recall test in Experiment 3. Performance is reported in percentages, and error bars report standard error of the mean.

Recognition results are reported in Figure 9. A 2 (verbal mediator vs. spontaneous mediator) X 2 (cue recognition or target recognition) mixed factor ANOVA revealed a main effect of recognition test format (with target recognition greater than cue recognition) \( F(1,34) = 23.33, MSE = 49.04, p < .001, \eta^2_p = .41 \), but no main effect of mediator group \( F(1,34) = 1.69, MSE = 274.11, p = .203, \eta^2_p = .05 \). Furthermore, the
interaction between recognition test format and mediator group was not significant \( F < 1 \). These cued recall and recognition results suggest that the overall performance advantage observed in forward versus backward cued recall are unlikely to be entirely attributed to the performance advantage observed in target versus cue accessibility (12 versus 8%, respectively), suggesting associative asymmetry. Consistent with the cued recall outcomes, participants recognized more target words versus cue words regardless of whether or not they were instructed to explicitly generate verbal mediators during practice. Together, these cued recall and recognition results suggest that verbal mediator use may not entirely explain the observed patterns of associative asymmetry.
FIGURE 9. Mean final recognition performance as a function of mediator group and type of cued recall test in Experiment 3. Performance is reported in percentages, and error bars report standard error of the mean.

Thus far, results from Experiments 2 and 3 provide evidence to suggest that verbal mediator use does not necessarily account for the observed patterns of associative asymmetry. However, performance thus far has included all items (regardless of what specific mediator they used during practice). Even though more sentence mediators and keyword mediators were used in the verbal mediator group versus spontaneous group, a moderate percentage of other strategies (e.g., repetition) were used by both groups. Therefore, I conducted conditional analyses to more precisely examine the influence of verbal mediators versus visual mediators on final test performance. These conditional analyses reflect items for which a participant indicated that they used a verbal mediator
(i.e., a sentence mediator or a keyword mediator), a visual mediator (i.e., imagery), or some other kind of strategy (including repetition, some other kind of strategy, or no strategy at all). Also, given that there was no significant difference between the verbal mediator group versus the spontaneous group in both the cued recall and recognition tests, I collapsed the conditional analyses across this variable.

Cued recall outcomes conditionalized on mediator type are reported in Figure 10. Given that I made no a priori predictions regarding how using a non-verbal or non-visual strategy would influence the pattern of associative asymmetry, analyses below focus on items for which participants reported using either a verbal or visual mediator during practice. A 2 (forward versus backward cued recall) x 2 (verbal mediator versus visual mediator) mixed factor ANOVA revealed a significant main effect of final test format (with greater performance in the forward versus backward direction) \( F(1,104) = 8.03, MSE = 1276.38, p = .006, \eta^2_p = .07 \). Additionally, there was a main effect of type of mediator used, with greater cued recall performance for visual mediators versus verbal mediators \( F(1,104) = 4.18, MSE = 350.40, p = .043, \eta^2_p = .04 \). Importantly, there was no interaction between final test format (forward versus backward cued recall) x type of mediator used (verbal mediator versus visual mediator) \( F < 1 \). These results demonstrate a persistent advantage for forward versus backward cued recall regardless of whether a verbal or visual mediator was used during practice. This pattern is not consistent with my original prediction that the pattern of associative asymmetry would be more pronounced with verbal versus visual mediators. Overall, these results provide
evidence to suggest that mediator use may not be a viable explanation for the previously observed patterns of associative asymmetry.

![Figure 10](image)

**FIGURE 10.** Mean final cued-recall performance as a function of mediator strategy and type of cued recall test in Experiment 3. Performance is reported in percentages, and error bars report standard error of the mean.

Recognition outcomes conditionalized on mediator type are reported in Figure 11. A 2 (cue or target recognition) X 2 (verbal mediator versus visual mediator) within-participant ANOVA revealed a significant main effect of recognition (with greater performance for target recognition versus cue recognition) \[F(1,32) = 18.04, MSE = 178.20, p < .001, \eta^2_p = .36\]. There was no main effect of type of mediator used \[F(1,32)\]
Importantly, the interaction between recognition (cue recognition versus target recognition) X type of mediator used (verbal mediator versus visual mediator) was not significant [$F < 1$].

FIGURE 11. Mean final recognition performance as a function of mediator strategy and recognition stimulus in Experiment 3. Performance is reported in percentages, and error bars report standard error of the mean.

Summary

Results demonstrate that there was a consistent performance advantage in forward versus backward cued recall, but also in target versus cue recognition. Thus, the patterns of associative asymmetry were not as clear as in Experiment 1. However, the
numerically greater difference in forward versus backward cued recall than in target versus cue recognition at least suggests that cued recall differences are unlikely to be entirely attributable to differences in target versus cue memory (suggesting associative asymmetry). Interestingly, these patterns emerged regardless of whether participants indicated that they had used a verbal mediator or a visual mediator during initial learning. I had originally hypothesized that verbal mediators would exaggerate and visual mediators would attenuate the performance differences between forward versus backward cued recall; however, this did not occur. These findings suggest that mediator use during initial learning minimally affects the observed pattern of associative asymmetry favoring forward versus backward cued recall performance.
GENERAL DISCUSSION

Experiment 1 clearly demonstrated that the effects of criterion level on associative memory are asymmetric. First, greater gains across criterion level were observed for forward versus backward cued recall. Second, these greater gains in forward versus backward cued recall emerged regardless of whether cue or target accessibility was similar (i.e., recognition performance in the pre-exposure condition) or favored target accessibility (i.e., recognition performance in the no-pre-exposure condition). These results replicate earlier work demonstrating associative asymmetry (Vaughn & Rawson, resubmitted). Unexpectedly, Experiments 2 and 3 yielded somewhat weaker patterns of associative asymmetry. In both experiments, cued recall performance favored forward versus backward cued recall; however, recognition performance also favored target versus cue recognition. Of interest, the performance advantage in forward versus backward cued recall was numerically greater than the performance advantage in target versus cue recognition. Therefore, the greater gains in forward versus backward cued recall are unlikely to be completely attributable to differences in target and cue accessibility (suggesting associative asymmetry).

Of greatest interest for present purposes, to what extent does mediator use underlie the asymmetric effects of criterion level on associative memory? Results from Experiment 2 established that participants use mediators spontaneously, which is necessary to establish for mediator use to be a viable account of the previously observed patterns of associative asymmetry (given that prior work demonstrating asymmetric
associative memory gains across criterion level did not explicitly examine mediator use during practice; Vaughn & Rawson, resubmitted). Experiment 3 further evaluated the extent to which mediator use may account for these observed patterns of associative asymmetry by manipulating mediator use. I expected performance patterns to remain asymmetric (i.e., greater forward versus backward cued recall performance despite similar cue and target recognition performance) when participants used verbal mediators during practice (i.e., sentence mediators or keyword mediators). These outcomes were only partially supported, as performance did favor forward versus backward cued recall but also favored target versus cue recognition. I also expected performance patterns to become more symmetric (i.e., similar forward and backward cued recall performance as well as similar cue and target recognition performance) when participants used visual mediators during practice (i.e., imagery). These outcomes did not emerge, as performance patterns were consistent with those observed when verbal mediators were used during practice (i.e., greater forward versus backward cued recall performance as well as greater target versus cue recognition performance). Overall, these results suggest that kind of mediator (either verbal or visual) does not appear to moderate the observed pattern of associative asymmetry.

Why did kind of mediator (either verbal or visual) not appear to moderate the observed pattern of associative asymmetry? One potential explanation is that even though participants spontaneously generated mediators during initial learning, participants may not spontaneously activate those mediators when completing the final test. For instance, imagine a participant forms the sentence “the mosquito was on the pipe” to remember the
word pair “mosquito-pipe”. During learning, he/she may activate this mediator on each test trial (e.g., “mosquito-???”) and use the mediator to help him/her reach criterion. However, on the final test after a week delay, it is possible that he/she no longer needs the mediator in order to successfully recall that “pipe” goes with the word “mosquito”. Relatedly, he/she may only search for and attempt to use the original mediator if he/she struggles to recall the correct word initially (i.e., the item is not easily recallable, and thus he/she tries to think back to the mediator he/she had originally generated during initial learning). In this scenario, participants may have recalled or failed to recall their original mediator (and subsequently may or may not have recalled the correct target word during the final test). In other words, without explicitly measuring if or how participants used their originally generated mediator on the final test, it is difficult to know how often (or in what ways) participants actually used their generated mediators on the final test. Future research could address this question by probing participants retrospectively (i.e., after all items had been tested in the final session) about the strategies they used to recall / recognize each item. Furthermore, in the spirit of Pyc and Rawson (2010), one could also measure mediator retrieval by asking participants to recall the mediators they had originally generated during practice for each item. Measuring mediator decoding would be less feasible given that the sentence mediator (e.g., “the mosquito sat on the pipe”) would likely contain both the cue word (e.g., “mosquito”) and the target word (e.g., “pipe”). In contrast, measuring mediator decoding would be more feasible if participants were instructed to generate keyword mediators (e.g., “smoke”) instead of sentence mediators (e.g., “the mosquito sat on the pipe”). Measures of mediator retrieval and
mediator decoding would help to diagnose how often participants spontaneously activate mediators during the final test.

Assuming participants do spontaneously use their originally generated mediators on the final test (or, at least use them to some extent), what other factors may be responsible for the lack of final test differences between the verbal mediator group versus the spontaneous group? I had originally predicted that the pattern of associative asymmetry would attenuate with the use of visual mediators; however, this was not observed. One potential explanation for this is that items may have evoked a mixture of mediator strategies during practice. For instance, it is quite plausible that participants who generated a sentence mediator during practice (e.g., “the mosquito was on the pipe”) also experienced some degree of interactive imagery (especially considering that the sentence “the mosquito was on the pipe” is a somewhat bizarre sentence). Similarly, participants who generated an interactive image during practice may not have used that image exclusively during practice (e.g., participants may have also related the words via a sentence, or thought about the words both in terms of a sentence and an image).

Although the mediator assessment instructions were worded to encourage participants to indicate which strategy they had used most often during practice, it is entirely plausible that certain items afforded more than one strategy during practice or involved some combination of strategies. Future research could investigate this possibility by allowing participants to select more than one strategy for each particular item (if need be) and conducting subsequent analyses for items in which only a verbal mediator or only a visual mediator was used exclusively.
Relatedly, the present materials could potentially be altered to more strongly investigate how verbal and visual mediators influence the pattern of associative asymmetry. Presently, the materials consisted of unrelated word pairs (e.g., “mosquito-pipe”) that were purposefully assembled to have zero forward and zero backward associative strength. Ensuring that the items have zero forward and zero backward associative strength is of primary importance when investigating associative symmetry because the cued recall tests must not differ in any aspect other than direction of recall (i.e., “queen” must be an equally effective cue for “journal” as “journal” is for “queen”). Another potentially important factor (or, at least a factor which one could manipulate) is word imageability (i.e., how easily a particular item affords a visual image). For instance, word pairs which consist of items with high imageability may be more prone to symmetrical forward and backward cued recall (versus items with low imageability). Although I did not directly manipulate item imageability in the present experiments, future research may seek to address how item properties (e.g., imageability) influence the pattern of associative asymmetry.

The primary purpose of the present research was to investigate to what extent mediator use may be able to account for the previously observed patterns of associative asymmetry. The results from Experiment 3 suggest that differences in forward versus backward recall persisted regardless of whether a verbal or visual mediator is used during practice. If mediator type cannot explain the previously observed patterns of associative asymmetry, what other factors may be responsible? According to the *elaborative retrieval hypothesis* (ERH; see Carpenter, 2006, 2009), testing improves memory via
increased semantic activation. Importantly, testing triggers a broad range of increased semantic activation (e.g., recalling the word pair “basket-bread” may trigger “basket → eggs → flour → bread”). Furthermore, testing triggers this broad range of semantic activation implicitly and automatically (i.e., participants are not necessarily aware of the various words or concepts they have activated during a test trial for a particular item). Of interest, the semantically activated information becomes forged into the cue-target pathway, such that a subsequent retrieval attempt (e.g., “bread-???”) has a greater chance of being successful if more versus less semantic information was originally activated (as greater semantic activation would provide additional retrieval cues to elicit the correct target word). There are several important points to note here: First, ERH claims that this process occurs automatically during test trials. Therefore, even if participants reported using a visual mediator during practice, ERH would claim that testing still triggered semantic activation, and that increased semantic activation is what is responsible for the benefits of testing. Second, it is important to notice that the semantically activated information begins with information related to the cue word (e.g., “bread”) and continues until some connection is made with the target word (e.g., “eggs”). In other words, the semantically activated information is asymmetric (i.e., favoring the cue → target direction). If testing benefits memory due to the increased activation of semantically related information, and if such activation occurs implicitly and favors the cue → target direction, then ERH provides a potential explanation for the observed asymmetry with visual mediators. Although speculative, final test performance may have been driven more by the implicitly activated semantic information rather than the explicitly generated
(or explicitly reported) mediators during practice (even for items in which participants reported using a visual mediator during practice).

In sum, the present work replicated prior research by showing that associative memory gains across criterion level appear to be asymmetric (Vaughn & Rawson, resubmitted). Results also suggested that even though participants spontaneously use mediators during practice, mediator use does not necessarily explain the observed patterns of associative asymmetry. Associative asymmetry was observed regardless of whether participants used a verbal mediator or a visual mediator during practice. Therefore, these results suggest that mediator use does not seem to provide a viable explanation for the previously observed patterns of associative asymmetry.
ENDNOTE

1. For purposes of completeness, there was no difference between forward versus backward cued recall performance as well as no difference between target versus cue recognition for items in the “other” category (i.e., items in which participants reported using a repetition strategy, no strategy, or some other kind of strategy representing a non-verbal or non-visual mediator) (all $F$s $< 2.02$).
REFERENCES


