

# Evaluating the Contribution of Emotional Valence to Associative Memory: Retrieval Practice Matters

Aiqing Nie, Guimei Jiang, and Mengmeng Li

Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, China

## ABSTRACT

Research has indicated that emotional valence can influence associative memory, but it is less clear whether it still works when the retrieval practice is controlled. The current study combined an associative recognition task with a paradigm of retrieval practice, with negative, neutral, and positive word pairs serving as stimuli. Results revealed that intact pairs possessed higher correct response proportions than rearranged, old+new, and new pairs; the rearranged pairs were more likely to be classified as intact; a negative impairment effect was observed in both learning conditions; the retrieval practice effect was sensitive to the interaction of emotional valence by pair type. We shows that the involvement of the recollection-driven process varies with pair type, providing telling evidence for the dual-process models; the occurrence of negative impairment effect conforms to the account of spontaneous interactive imagery; the contribution of desirable difficulty framework is modulated by the interaction of emotional valence by pair type.

## KEYWORDS

associative memory  
emotional valence  
retrieval practice  
pairs  
dual-process models

## INTRODUCTION

### Associative Memory and the Dual-Process Models

In daily life, one may undergo numerous events, and some of them are bound together. For instance, one may have a hamburger and a glass of milk for breakfast, and afterwards, he/she may remember eating a hamburger but forget about the milk, or remember drinking the milk but forget the hamburger, or remember/forget both. In the memory domain, the capability to remember events experienced together, such as two pieces of information, is operationally defined as *associative memory* (Baadte & Meinhardt-Injac, 2019; Bellander et al., 2017; Leach et al., 2019; Liu et al., 2021; Lucas et al., 2019; Maylor et al., 2019; Nie

& Jiang, 2019; Osth & Fox, 2019). To explore how the bound events are memorized, researchers have designed a typical associative recognition paradigm in which several strings of pairs (e.g., A-B, C-D, and E-F) are studied, and in a following test, participants are instructed to indicate whether a pair is intact (exactly the same as the one in preceding study phase, e.g., A-B), rearranged (both items are studied but are recombined, e.g., C-F), or new (both items in it are novel, e.g., G-H; Bellander et al., 2017; Bridger et al., 2017; De Brigard et al., 2020; Delhay et al., 2019; Leach et al., 2019; Li et al., 2019; Maylor et al., 2019; Osth & Fox,

Corresponding author: Aiqing Nie, Department of Psychology and Behavioral Sciences, Zhejiang University, 148 Tianmushan Road, Hangzhou, China 310028.  
E-mail: nieaiq@126.com

2019). Studies with this paradigm generally demonstrate that the identifications of both intact and new pairs are much easier than those of rearranged ones (Bridger et al., 2017; Li et al., 2019).

Besides the typical associative recognition paradigm, several variants have been designed, such as the *selective recognition paradigm* that requires participants to identify the formerly associated item for a given cue from several alternative options (Caplan et al., 2019; Huguet et al., 2019), as well as the *cued recall paradigm* in which participants recall the previously associated item for a cue (Madan et al., 2019; Siegel et al., 2019). In addition, to further disentangle the subtle difference of associative information, a fourth pair, the *old+new* pair, is also introduced, in which one item is originated from a studied pair while the other item is not experienced before (e.g., D-I; Buchler et al., 2011; Buchler et al., 2008; Nie & Jiang, 2019). Extant investigations have demonstrated that the memory performance of old+new pairs is significantly different from those of intact, rearranged, and new pairs, showing that the discriminations of old+new pairs are much easier than those of rearranged ones, but much harder than those of intact and new ones (Buchler et al., 2008, 2011; Nie & Jiang, 2019).

Dual-process models claim that memory hinges on two independent processes: familiarity and recollection (Li et al., 2019; Liu et al., 2021; Maylor et al., 2019; Nie & Jiang, 2019; Nie et al., 2019; Ye et al., 2019; Zhou et al., 2020). Familiarity is an automatic and fast-acting process, involving a generic feeling that something/someone has been previously encountered. However, recollection is a slow and controlled process with autonoetic consciousness that involves planning and monitoring that enables one to reinstate explicit contextual details associated with past encounters, such as when or where an item was experienced. For associative memory, the engagements of these two processes vary with the pair type. Buchler et al. (2011) have suggested that the discriminations of intact and rearranged pairs are primarily dependent upon the recollection-based process, since participants should determine whether the two items within a pair originally occur simultaneously or not. In that case, if only the involvement of the familiarity-based process but not the recollection-based process supports the recognition of the items in the studied pairs, it may fail to facilitate the identification between intact and rearranged pairs. In contrast, the discernments of old+new and new pairs may rely less on recollection-based process, as a feeling of familiarity regarding whether an item has been previously experienced is enough to support such discernments.

## Negative Effect and Positive Bias in Associative Memory, and Possible Accounts

So far, a series of influencing factors have been confirmed for associative memory, including the emotional valence of stimuli, which is one of our interests. There is debate over the contribution of emotional valence of stimuli to associative memory, since the findings are mixed (Caplan et al., 2019; Madan et al., 2017, 2019; Nadarevic, 2017; Siegel et al., 2019; Zimmerman & Kelley, 2010). Some investigations revealed that the negative valence of stimuli could impair associative memory. For example, Madan et al. (2017) and Caplan et al. (2019) both veri-

fied a *negative impairment effect*, in which pure negative pairs (i.e., pairs of *negative item-negative item*) were more poorly identified than pure neutral ones (i.e., pairs of *neutral item-neutral item*). This could be accounted for from different perspectives. First, it was explained by the lower effectiveness of the spontaneous *interactive imagery* of negative pairs. Interactive imagery means forming a mental image that combines elements of an association, and it is known to trigger the unitization of the elements and occur spontaneously (Ahmad & Hockley, 2014; Caplan et al., 2019). Extant investigations have demonstrated that, compared with neutral pairs, spontaneous interactive imagery is relatively poorer in negative cases when forming between-item associations to operate the items as a single one, which renders the encoding for negative pairs worse (Caplan et al., 2019). Second, the judgment of learning (JoL) plays an important role, since knowing the memorability of an experience is particularly important for memory control processes, such as investing time and effort in encoding when one wants to guarantee that they will remember. To be specific, JoL means the judgments made by participants at the end of a learning trial regarding the likelihood of remembering the acquired information on a subsequent memory test. In other words, it is the participant's subjective feeling about whether they will remember the learned information in the future (Koriat, 1997; Zimmerman & Kelley, 2010). In addition, a previous study found that the lower-JoL of neutral pairs tended to facilitate more study efforts while the higher-JoL of negative pairs would facilitate less study efforts, resulting in a memory reduction for negative pairs (Zimmerman & Kelley, 2010). Lastly, attentional cost might also be a contributor: For instance, attentional accounts propose that emotional aspects of a scene capture attentional resources to enhance memory for the emotional item at the expense of inhibiting nonemotional peripheral information, making the increased attention toward an emotional stimulus plays a role in reducing memory for associations (e.g., Bisby & Burgess, 2014).

By contrast, other studies confirmed the opposite results. For instance, the pairs of *nonword-negative word* performed better than the pairs of *nonword-neutral word* (Nadarevic, 2017). Also, a study using sentences revealed greater performance in negative conditions (Siegel et al., 2019), manifesting as the negative improvement effect. The binding hypothesis offers a supporting explanation for these findings, as it asserts that the emotionally aroused stimuli hold more binding superiority than neutral ones (Nadarevic, 2017; Siegel et al., 2019). Additionally, Siegel et al. also discovered a memory benefit for negative rather than positive sentences, suggesting that more processing resources and detail-oriented operations were assigned to negative cases.

Furthermore, some investigations have verified a *positive bias* in associative memory. Zimmerman and Kelley (2010) found that compared with both pure negative and pure neutral pairs, pure positive pairs (i.e., pairs of *positive item-positive item*) showed a superiority in associative memory, that is, a positive bias. Pure neutral ones behaved better than pure negative ones. Also, Madan et al. (2019) confirmed a positive bias, showing that pure positive pairs held better performance than all other types of word pairs, namely, the pure pairs of another valence (e.g., neutral item-neutral item), as well as the mixed pairs of

positive item-neutral item and neutral item-positive item. The positive bias could possibly be explained through the broaden-and-build theory, as it states that positive emotional valence of stimuli can attract wide scopes of attention, perception, and cognition, while negative emotional valence is hypothesized to narrow the field of attention, perception, and cognition, which subsequently enables one to build a holistic processing for the positive stimuli, thus resulting in the positive bias (Fredrickson, 2013; Madan et al., 2019; Schmitz et al., 2009; Zimmerman & Kelley, 2010). In brief, the above-mentioned research demonstrated distinct contributions of emotional valence of stimuli to associative memory, exhibiting either a negative impairment effect, a negative improvement effect, or a positive bias. After reviewing the literature, we summarized several latent causes for this discrepancy or complexity. The first justification referred to the different research paradigms, since Madan et al. (2017) and Caplan et al. (2019) adopted the selective recognition manner, while others applied the cued recall paradigm (Madan et al., 2019; Nadarevic, 2017; Siegel et al., 2019; Zimmerman & Kelley, 2010). Second, the emotional stimuli were varied: some studies considered negative, neutral, and positive pairs simultaneously (Siegel et al., 2019; Zimmerman & Kelley, 2010), while others only compared neutral pairs with another emotional valence, either negative pairs (Caplan et al., 2019; Madan et al., 2017; Nadarevic, 2017), or positive pairs (Madan et al., 2019). Third, the compositions of pairs were different: pure pairs (Caplan et al., 2019; Madan et al., 2017; Zimmerman & Kelley, 2010), or mixed pairs (Madan et al., 2019; Nadarevic, 2017; Siegel et al., 2019). Last, materials of diverse modality were applied, such as pictures (Caplan et al., 2019; Madan et al., 2017), words or nonwords (Madan et al., 2019; Siegel et al., 2019; Zimmerman & Kelley, 2010), and sentences (Nadarevic, 2017). Given these issues, we focused on emotional words, and our first objective was to further examine how the emotional valence of stimuli affects associative memory. For this reason, the associative recognition paradigm was utilized, and pure word pairs of three types of emotional valences (negative, neutral, and positive) were used as stimuli.

## Retrieval Practice Effect and Possible Accounts

In school education, testing is an effective way to figure out whether a student has mastered the studied materials well. If not, they should restudy the materials and engage in another test, and the improved test scores tend to confirm that restudy is an effective way to enhance performance. However, the test itself also matters. To examine which one is more effective, test or restudy, many researchers have adopted a three-phase paradigm in memory, which typically includes study, retrieval practice/restudy, and a final test. The participants are first instructed to study several stimulus pairs. Subsequently, some pairs are put into a cued recall task (i.e., retrieval practice condition), while others are assigned to restudy (i.e., restudy condition). Finally, all studied pairs include a final cued recall test (Abel & Roediger, 2017; Buchin & Mulligan, 2017; Szöllösi et al., 2017).

Interestingly, studies adopting this paradigm generally revealed that during the final test, the pairs from the retrieval practice condition

were better remembered than those from the restudy condition. This difference was termed the *retrieval practice effect* (Abel & Roediger, 2017; Binks, 2018; Buchin & Mulligan, 2017; Cho, Neely, Brennan et al., 2017; Cho, Neely, Crocco et al., 2017; Peng et al., 2019; Peterson & Wissman, 2018; Rickard & Pan, 2018; Szöllösi et al., 2017). Moreover, even when the manners of retrieval practice or final test in the typical paradigm are altered discretionarily, significant retrieval practice effects can still be observed. For instance, there are studies where both retrieval practice and the final test were free recall tasks (Bae et al., 2018; Beardsley et al., 2018), free recall for retrieval practice but a recognition task for the final test (Huff et al., 2018; Verkoeijen et al., 2012), cued recall for retrieval practice but a recognition task for the final test (Liu et al., 2017), or cued recall for retrieval practice but a free recall task for the final test (Mulligan et al., 2016). It follows that the retrieval practice effect is seemingly not shaped by the different manners of retrieval practice and of final test.

Thus far, the desirable difficulty framework offers reasonable explanations for the retrieval practice effect. It claims that, compared with restudy, retrieval practice demands more effort; this effort makes the to-be remembered events in the retrieval practice condition more retentive, and eventually results in enhanced availability and accessibility in the final test (Binks, 2018; Greving & Richter, 2018; Kubik et al., 2018; Rickard & Pan, 2018). It is certainly worth noting that retrievability plays an essential role in this framework. When considering the correctly retrieved information, the performance would be much higher in the retrieval practice than in the restudy condition, demonstrating the anticipated retrieval practice effect.

For all we know, many potential influence factors have been examined in the retrieval practice effect, such as feedback. If feedback is given, a larger retrieval practice effect emerges, which shows the facilitation effect of feedback (Minear et al., 2018). However, extant literature on the retrieval practice effect mainly concerns neutral stimuli, while data on emotional ones are comparatively scarce. Existing studies have not found the contribution of the emotional valence of stimuli to the retrieval practice effect. For instance, Emmerdinger and Kuhbandner (2018) revealed no retrieval practice effects for any of the negative, neutral, and positive stimuli. Additionally, other studies found that the retrieval practice effect was insensitive to emotional valence of the stimuli (Emmerdinger et al., 2018; Jia et al., 2018). Thus, when the three-phase paradigm of retrieval practice effect is introduced into the memory domain, the emotional valence of the stimuli does not seem to work as an influence factor.

However, to the best of our knowledge, the extant literature on the retrieval practice effect merely focused on the impact of emotional valence on memory, while no study of three-phase paradigm has taken the influence of emotional valence of stimuli on associative memory into consideration. This issue, combined with our first objective, led us to examine whether the contribution of emotional valence of stimuli to associative memory is sensitive to retrieval practice. To this end, we introduced the three-phase paradigm of retrieval practice into associative memory, and also set feedback in the second phase.

## The Current Study and our Hypotheses

In sum, prior investigations have demonstrated reliable contributions of emotional valence of stimuli to associative memory, but the patterns are not always consistent, showing either a negative superiority/impairment or a positive superiority (i.e., negative effect or positive bias). When a phase of retrieval practice/restudy is inserted between the first study phase and the final test phase, reliable retrieval practice effect is verified. However, it remains unknown whether the contribution of emotional valence of stimuli to associative memory is invariable when retrieval practice/restudy is taken into consideration. To address this issue, the current study used the three-phase paradigm of retrieval practice, aiming to further examine the influence of emotional valence of stimuli on associative memory. Pure word pairs of different emotional valences (i.e., negative, neutral, and positive) served as experimental materials, and the test pairs included intact, rearranged, old+new, and new ones. In addition, to enhance the role of retrieval practice, feedback was arranged in the second phase.

First, we predicted that the performance (i.e., correct response proportions, old proportions, and RTs) would behave as the function of pair type, such that the correct response proportions would be the highest for intact pairs, followed by new pairs, and then by old+new pairs, while the proportions would be the lowest for rearranged pairs. The best performance for intact pairs might be produced by the manipulation of the second phase. To assess how much more likely it would be for participants to classify both rearranged and old+new pairs as intact ones than as new pairs, we carried out a further analysis for old proportions (i.e., response proportions for different pairs as intact ones). We predicted that the old proportions would also exhibit similar patterns, that is, the proportions of intact pairs would be the largest, followed by the new and old+new pairs in succession, while the rearranged pairs would be the last. Regarding RTs, the discrimination of intact pairs would be the fastest, followed by new pairs, and then by old+new pairs, while rearranged pairs would be the slowest.

Second, it was hypothesized that the performance would be sensitive to the factor of emotional valence of stimuli. However, it was difficult to clarify whether there would be a benefit of emotional valence, as previous studies revealed mixed patterns, with either a facilitation or inhibition of emotional stimuli (including either negative or positive stimuli, or both) to memory performance (see above). Based on this, we anticipated either a negative effect or a positive bias. Specifically, if the performance of negative pairs was inferior to that of neutral pairs, a reliable negative impairment effect would be shown, and the spontaneous interactive imagery would be supported. If the opposite pattern was observed, a negative improvement effect would be confirmed, seeing that negative pairs would occupy more processing resources and be processed in a detail-oriented manner. If positive pairs had a better performance than neutral and negative pairs, a positive bias would be revealed, and thus the broaden-and-build theory would be proven. If negative and positive pairs showed higher performance than neutral pairs, both the negative improvement effect and positive bias would be observed, and thus the binding hypothesis would be proven.

In addition, we hypothesized that the learning condition (retrieval practice/restudy) would interact with emotional valence, showing that the hypothetically negative effect or positive bias would be weakened or show different patterns. We attributed these to the actively conscious process (i.e., effort) from retrieval practice and to the exposure from the restudy, which might collide with the unique processing mechanism of emotional stimuli, resulting in the inefficiency of the negative effect or positive bias.

In addition, we assumed that emotional valence would interact with pair type, such that in the intact pair case, the performance of neutral pairs would be superior to that of negative ones but inferior to that of positive ones; whereas there would be neither emotional valence difference nor different patterns for rearranged, old+new, and new pairs. This emotional valence difference among different pair types might originate from the different extents of the involvement of familiarity- and recollection-based processes.

Third, we expected significant retrieval practice effect, which would support the desirable difficulty framework, since more effort would be made under the retrieval practice condition so the pairs experienced a retrieval practice would have a higher availability and accessibility than those under the restudy condition. Furthermore, the retrieval practice effect was expected to be susceptible to the emotional valence of stimuli, demonstrating that negative pairs benefit less from retrieval practice than neutral and positive ones. On the other hand, it might be insusceptible to the emotional valence of words, showing that the retrieval practice effect is not only present across the three types of emotional valence but also shares a similar benefit effect among these conditions.

Last, the retrieval practice effect would be modulated by the emotional valence of stimuli and pair type, such that negative or positive intact pairs would benefit more from retrieval practice than neutral intact ones, while the retrieval practice benefit of negative or positive rearranged and old+new pairs would exhibit an opposite pattern. These discrepancies could be attributed to the interaction between the unique processing mechanism of emotional stimuli and the unequal involvement of recollection-based process across different pairs.

## METHODS

### Participants

A total of 34 undergraduate students (26 females) were recruited. Their age ranged from 18 to 22 years with a mean of  $20.324 \pm 0.629$  years. Four were excluded from data analyses, two owing to the fact that they did not comply with the task instructions, and two due to the fact that their data did not conform to the standard deviation principle. All participants spoke Chinese as their first language and were right-handed. All had normal or corrected-to-normal visual acuity, and none of them reported any problem of neurological impairments, learning disabilities, or psychiatric disorders. All participants signed a written informed consent before participating in the experiment, and received a compensation of course credits for voluntary participation. The cur-

rent protocols were approved by the Research Ethics Committee of Zhejiang University, and also met the ethical standards laid out by the Declaration of Helsinki. When the experiment was completed, each participant was debriefed and thanked.

Considering the sample size, to assess how many participants should be recruited, an *a priori* power analysis using the *G\*power* 3.1 software (Faul et al., 2009) was run for the current variables. It revealed that if an *f* effect size of 0.25 and the power of 0.80 with an  $\alpha$  restriction of .05 were expected, a minimal sample size of 28 was necessary. Thus, the final sample size of 30 was appropriate.

## Design

This experiment was a 4 (pair type: intact, rearranged, old+new, and new)  $\times$  2 (learning condition: retrieval practice and restudy)  $\times$  3 (emotional valence: negative, neutral, and positive) incomplete within-subject design. Pair type referred to the status of test pairs, the two levels of learning condition were manipulated by the second phase of three-phase paradigm, and emotional valence was that of pairs (i.e., the valence of both words in each pair). All pairs were pure emotional ones, in line with previous studies (Caplan et al., 2019; Madan et al., 2017; Nie & Jiang, 2019). The three levels of intact, rearranged, old+new of pair type all had the two levels of learning condition, and also the three levels of emotional valence, and such a pattern held true for learning condition and emotional valence; if only these levels were considered, our design would be a complete within-subject design. However, we had the new level for pair type in which the items in each new pair were all novel or not experienced previously, such that this type of pair experienced neither a state of retrieval practice nor restudy, and thus made the current experiment an incomplete within-subject design.

## Materials

The current study used 912 words as stimuli, originally acquired from the Affective Norms for English Words (ANEW, Bradley & Lang, 1999), all translated into Chinese as in one of our recent studies (Nie & Jiang, 2019). Among them, 840 words served as formal materials. The emotional valence of the words was classified on the basis of the standard in Berger et al. (2018), that is, valences less than 3.5 were classified as negative, 3.5 to 6.5 as neutral, and greater than 6.5 as positive. Thus, for the formal stimuli, there were two hundred and eighty words per emotional valence. The words within each valence were pseudo-randomly paired, resulting in 140 pairs. Besides the formal stimuli, the remaining 72 words (i.e., 36 pairs) served as practice trials (i.e., 16) and fillers (i.e., 20 pairs), and they were also equal in numbers across the three types of emotional valence.

To avoid interference across distinct pairs and reduce participants' fatigue, the formal stimuli were evenly subdivided into 10 blocks, forming 42 pairs per block, with 14 pairs per emotional valence. The words and the word pairs were all different across all these 10 blocks, that is, the stimuli were not repeated in different blocks. Five blocks were assigned to the retrieval practice condition while the other five blocks were allocated to the restudy condition. Within each block, there were 30 pairs in the study phase, that is, 10 pairs per emotional valence. The

pairs in the following retrieval practice/restudy course were identical to those in the study session. There were 42 pairs (i.e., 14 pairs per emotional valence) in the final test, which contained 12 intact, 12 rearranged, 12 old+new, and six new pairs, and these four types were also controlled for emotional valence. Thus, every participant experienced all the words and the word pairs.

During the final test phase, the intact pairs were completely identical to those in the two previous phases, and their corresponding learning conditions were determined by their status in the second phase. Each rearranged pair had two studied words which previously belonged to different pairs, and both words had the same history of learning condition, either of retrieval practice or restudy. There was merely one studied word per old+new pair, that is, the pairs were formed by randomly pairing the words of six studied pairs with 12 unstudied words, and the learning condition was determined by that of the studied word in the previous phase within each pair. As stated above, the two words in each new pair were both novel, making such a pair have no state of learning condition. In order to eliminate primacy and recency effects, each block had two pairs as fillers separately displayed at the beginning and the end of the study phase. Fillers were not involved in the following test session.

## Procedure

During the experiment, participants comfortably sat in a quiet room with dim lighting. They were instructed to engage in the practice trials plus ten formal blocks. All protocols were identical between the practice and the formal experiment, and the task instructions were separately presented and described to each participant before each phase. Before the formal blocks, each participant was given the practice trials of three phases to familiarize them with the experimental protocols. Below, we describe each phase separately, along with the distraction. In addition, due to the particularity of the second phase, the two learning conditions, retrieval practice and restudy sessions, are described separately.

### STUDY PHASE

Within the study session, the instructions were first given to each participant, requiring them to study upcoming pairs through reading with no other requirements emphasized. Then, trials of a fixation cross and a word pair began. Within each trial, a fixation cross was first displayed in the center of the computer monitor for 500 ms, followed by a word pair lasting for 3 s. Then the next trial was displayed.

### DISTRACTION

When each study session was finished, there was a distraction task of three-digit backward counting. The purpose of the distraction was to remove all word information from working memory. Considering the distraction task, a three-digit Arabic number (e.g., 763) was presented in the center of the screen for 1 min. The participants continuously subtracted three from the number. They orally reported each answer until the original number disappeared.



## RETRIEVAL PRACTICE SESSION

In this session, the instructions were first described to the participants. They were informed that different cues (i.e., one word per studied pair) would be showed on the screen and they should retrieve its word paired during the prior study phase, and then press the *S* or *L* key on the keyboard to indicate whether they could remember the associated word or not. The keys were pressed by both index fingers, and the assignments of response keys and response fingers were counter-balanced among blocks and among participants. Afterwards, a fixation cross was centrally displayed on the screen for 500 ms, followed by the presentation of one cue plus an empty underline for 3 s, during which a key press was made. The cues were displayed on the left side of the screen, while the underlines were shown on the right side, ensuring a 4 cm blank distance between them. When the 3 s ended, feedback of the cue and its previously paired word would be displayed for 2 s. Feedback was shown regardless of whether the participants had pressed the keys or not, or which key had been pressed. Then, the next trial proceeded (including the fixation cross, the presentation of cue and underline, as well as the following feedback). To ensure the pairs with identical emotional valence would not continuously occur more than three times, trials were displayed pseudo-randomly.

## RESTDY SESSION

In each restudy session, instructions were first described to each participant, informing them to restudy the pairs that had been previously studied. Next, the trials of a fixation cross and a word pair were presented successively. Within each trial, the fixation cross was centrally displayed for 500 ms, followed by a word pair displayed for 5 s.

## DISTRACTION

When each retrieval practice/restudy session was finished, a distraction task followed. Except the original number, other manipulations were the same as those in the distraction task after each study session.

## FINAL TEST

Before the emergence of the test pairs, the participants were instructed that several word pairs would be presented, and they should classify the pairs as intact, rearranged, old+new, or new ones, by pressing *D*, *F*, *J*, or *K* on the keyboard, respectively. The keys were pressed by index and middle fingers of both hands, and the assignments of response keys, response fingers, and response hands were counterbalanced among blocks and among participants. Afterwards, the trials of a fixation cross and the following pair were shown. In one trial, the fixation cross was presented in the center of the screen for 500 ms and the following pair for 4 s. The presentation of the next trial proceeded regardless of whether a response had been made, or the accuracy of response.

## OTHER MANIPULATIONS

The experiment was programmed in *E-prime* software 2.0 (Psychology Software Tools, Inc.), and was run on a laboratory computer equipped with a 17 in. LCD monitor. The monitor resolution was  $1.024 \times 768$  px, and the refresh frequency was 100 Hz. The viewing

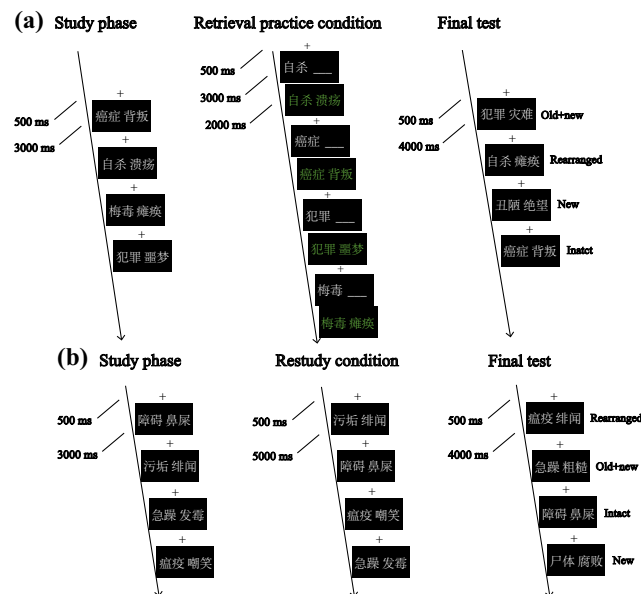
distance of stimuli was approximately 70 cm. The feedback of the cue and its paired word were displayed in green, while the instructions, fixation cross, numbers, cue and empty underline, and word pairs were all in white. The reason for displaying the feedback in green was to help participants track which session they were engaging in, in case that they might confound the tasks if the stimuli of all sessions were presented in white. All were against a black background. Instructions were presented in the standard Chinese font size of No. 50. Fixation cross and numbers were in the font size of No. 100, and pairs were in No. 80 size. Considering the pairs during the study, restudy, and final test phases, one word was shown on the left side screen while the other was shown on the right side. The distance between the two words of each pair was always 4 cm.

The manipulations of both the study session and the final test were identical to the associative recognition paradigm, except the retrieval practice/restudy phase. For the sake of avoiding interference between different learning conditions and self-tests by participants, the five blocks of retrieval practice condition and those of the restudy condition were performed separately, in line with previous studies (Verkoeijen et al., 2012). To ensure the pairs with identical emotional valence would never continuously occur for more than three times, the pairs were pseudo-randomly displayed in each phase. Likewise, the presentations of the four pair types in the final test were pseudo-random as well. Participants were instructed to fixate on the center of the screen until the stimuli were displayed, and were encouraged to respond as quickly and accurately as possible. In order to prevent interferences across two consecutive blocks and to relieve the participants' fatigue, when an entire block was completed, there was always a 5 min interval for rest. The entire experiment lasted about two and a half hours. Some negative stimuli examples and a schematic procedure are illustrated in Figure 1. It should be noted that the word pairs of different emotional valences were pseudo-randomly displayed in all phases for each block. We used negative words as examples in Figure 1 to exhibit different pair types during the tests.

## STATISTICAL ANALYSES AND RESULTS

The statistical analysis did not include the data whose SDs were greater or lesser than 3. IBM SPSS Statistics v22 (IBM Corporation, 2014) was used to analyze the data. In case of sphericity violation, the Greenhouse-Geisser correction was adopted. *F* ratios are given with Greenhouse-Geisser adjusted *p* values, effect size of partial eta-squared ( $\eta_p^2$ ), and epsilon values ( $\epsilon$ ). Whenever necessary, the Bonferroni correction was used to counteract the problem of multiple comparisons where the 95% CIs were also reported. An  $\alpha$  level of .05 (two-tailed) was applied to all inferential analyses.

In the current study, the participants were required to press keys in both retrieval practice session and the final test phase. For the retrieval practice session, the repeated-measures analysis of variance (ANOVA) revealed no reliable main effect of emotional valence,  $F(2, 58) = 1.233, p = .293, \eta_p^2 = .041, \epsilon = 0.766$ , which showed that the recall performance did not significantly differ across negative, neutral, and

**FIGURE 1.**

Negative stimuli examples and a schematic procedure. In the study phase, participants were presented with a series of word pairs. In the retrieval practice session, participants retrieved the paired words for the cues (i.e., one word of per studied pair); in the restudy session, participants studied the pairs again. The final test was to differentiate among the four types of pairs: intact, rearranged, old+new, and new.

positive pairs ( $M_{\text{Negative}} = 0.424$ ,  $SE = 0.030$ ,  $M_{\text{Neutral}} = 0.449$ ,  $SE = 0.034$ , and  $M_{\text{Positive}} = 0.415$ ,  $SE = 0.030$ ). Hence, we did not analyze the results of this phase further. Next, to explore whether our assumptions were correct, data analysis was performed for correct response proportions, old proportions, and RTs separately, focusing only on the data from the final test phase. Unless otherwise specified, the data from both retrieval practice and restudy conditions were all of the final test phase. Only significant results were reported.

## Analyses for Correct Response Proportions and the Results

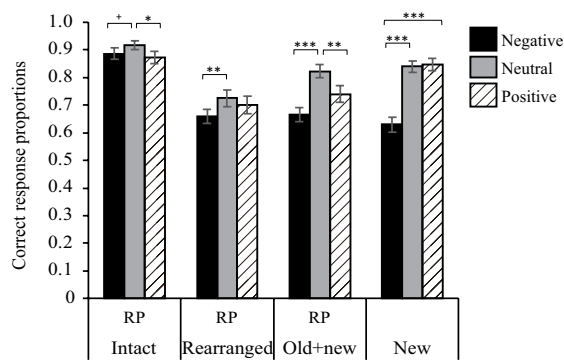
First, in order to test whether the correct response proportions behaved as a function of pair type by emotional valence, the proportions were submitted to a 4 (pair type: intact, rearranged, old+new, and new)  $\times$  3 (emotional valence: negative, neutral, and positive) repeated-measures ANOVA for retrieval practice condition and restudy condition separately. The reason for conducting a separate analysis was that this experiment was an incomplete within-subject design, where intact, rearranged, and old+new pairs all experienced either the retrieval practice condition or the restudy condition, but the new pairs did not, such that it was necessary to differentiate between two learning conditions to guarantee equality across different pair types, ensuring a rigorous exploration for the influence of emotional valence on associative memory.

Second, to substantiate whether the correct response proportions were sensitive to the learning condition and whether a reliable retrieval practice effect occurred, the proportions were subjected to a 3 (pair type: intact, rearranged, and old+new)  $\times$  2 (learning condition: retrieval practice and restudy)  $\times$  3 (emotional valence: negative, neutral, and positive) repeated-measures ANOVA.

## Analyses for Correct Response Proportions in Retrieval Practice Condition

The correct response proportions as the function of pair type by emotional valence in retrieval practice condition are illustrated in Figure 2. For this condition, the repeated-measures ANOVA of pair type by emotional valence revealed a significant main effect of pair type,  $F(3, 87) = 24.077$ ,  $p < .001$ ,  $\eta_p^2 = .454$ ,  $\epsilon = 0.565$ , which demonstrated that the proportions were much higher for intact pairs than for rearranged, old+new, and new pairs ( $p < .001$ , 95% CI [0.146, 0.249];  $p < .001$ , 95% CI [0.105, 0.194];  $p = .001$ , 95% CI [0.042, 0.198]). There was a significant main effect of emotional valence,  $F(2, 58) = 57.963$ ,  $p < .001$ ,  $\eta_p^2 = .667$ ,  $\epsilon = 0.948$ , showing that neutral pairs had much higher proportions than both negative and positive pairs ( $p < .001$ , 95% CI [0.088, 0.143];  $p = .003$ , 95% CI [0.011, 0.062]), and proportions for positive pairs were much higher than negative pairs ( $p < .001$ , 95% CI [0.049, 0.110]).

Additionally, the interaction between pair type and emotional valence was significant,  $F(6, 174) = 13.179$ ,  $p < .001$ ,  $\eta_p^2 = .312$ ,  $\epsilon = 0.751$ . Simple effect tests for this interaction found that, in the negative case, the correct response proportions were much higher for intact pairs than for the other three types of pairs ( $p < .001$ , 95% CI [0.173, 0.280];  $p < .001$ , 95% CI [0.143, 0.300];  $p < .001$ , 95% CI [0.156, 0.358]). In the neutral case, intact pairs had much higher proportions than the other three types of pairs ( $p < .001$ , 95% CI [0.121, 0.263];  $p < .001$ , 95% CI [0.048, 0.138];  $p = .030$ , 95% CI [0.005, 0.148]), and proportions for both old+new and new pairs were much higher than rearranged pairs ( $p = .003$ , 95% CI [0.026, 0.170];  $p = .028$ , 95% CI [0.009, 0.221]). In the positive case, intact pairs had much higher proportions than both rear-

**FIGURE 2.**

Correct response proportions as the function of pair type by emotional valence in the retrieval practice condition. RP = retrieval practice. Bars represent the SEM.

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

ranged and old+new pairs ( $p < .001$ , 95% CI [0.098, 0.249];  $p < .001$ , 95% CI [0.060, 0.206]), and proportions for new pairs were higher than both rearranged and old+new pairs ( $p = .008$ , 95% CI [0.030, 0.264];  $p = .030$ , 95% CI [0.007, 0.206]).

Regarding the factor of emotional valence in the two-way interaction, simple effect tests confirmed that, in the intact condition, neutral pairs had much higher proportions than both negative and positive ones ( $p = .052$ , 95% CI [0.000, 0.060];  $p = .037$ , 95% CI [0.002, 0.085]). In the rearranged condition, the proportions were much higher for neutral pairs than for negative ones ( $p = .006$ , 95% CI [0.016, 0.114]). Regarding the old+new condition, neutral pairs held much higher proportions than both negative and positive ones ( $p < .001$ , 95% CI [0.089, 0.227];  $p = .006$ , 95% CI [0.021, 0.146]). In the new condition, the proportions were much higher for both neutral and positive pairs than for negative ones ( $p < .001$ , 95% CI [0.160, 0.260];  $p < .001$ , 95% CI [0.162, 0.271]).

In brief, in terms of the retrieval practice condition, intact pairs had greater correct response proportions than rearranged, old+new, and new pairs, which indicated that the identifications of intact pairs were relatively easier. The correct response proportions were much higher for neutral than for positive pairs, which were followed by negative pairs, demonstrating a negative impairment effect. The correct response proportions were modulated by the interaction of pair type by emotional valence.

## Analyses for Correct Response Proportions in the Restudy Condition

For the restudy condition, the correct response proportions as the function of pair type by emotional valence are illustrated in Figure 3. In this condition, the repeated-measures ANOVA of pair type by emotional valence confirmed a reliable main effect of pair type,  $F(3, 87) = 8.913$ ,  $p < .001$ ,  $\eta_p^2 = .235$ ,  $\epsilon = 0.694$ , which indicated that the proportions were much higher for intact pairs than for both rearranged and old+new pairs ( $p < .001$ , 95% CI [0.052, 0.185];  $p < .001$ , 95% CI [0.071,

0.192]). There was also a reliable main effect of emotional valence,  $F(2, 58) = 38.331$ ,  $p < .001$ ,  $\eta_p^2 = .569$ ,  $\epsilon = 0.942$ , showing that neutral pairs had much higher proportions than both negative and positive pairs ( $p < .001$ , 95% CI [0.063, 0.122];  $p < .001$ , 95% CI [0.029, 0.084]), and proportions for positive pairs were higher than negative pairs ( $p = .002$ , 95% CI [0.012, 0.060]).

Moreover, the interaction of pair type by emotional valence reached statistical significance as well,  $F(6, 174) = 17.095$ ,  $p < .001$ ,  $\eta_p^2 = .371$ ,  $\epsilon = 0.749$ . Simple effect tests revealed that, in the negative condition, intact pairs had much higher correct response proportions than the other three types of pairs ( $p = .035$ , 95% CI [0.005, 0.212];  $p < .001$ , 95% CI [0.079, 0.221];  $p < .001$ , 95% CI [0.100, 0.327]). In the neutral condition, the proportions of intact pairs were much higher than those of both rearranged and old+new pairs ( $p = .007$ , 95% CI [0.020, 0.167];  $p = .007$ , 95% CI [0.019, 0.161]). In the positive condition, the proportions were much higher for intact pairs than for both rearranged and old+new pairs ( $p < .001$ , 95% CI [0.075, 0.235];  $p < .001$ , 95% CI [0.070, 0.240]), and new pairs also had significantly higher proportions than both rearranged and old+new pairs ( $p = .001$ , 95% CI [0.059, 0.271];  $p < .001$ , 95% CI [0.075, 0.255]).

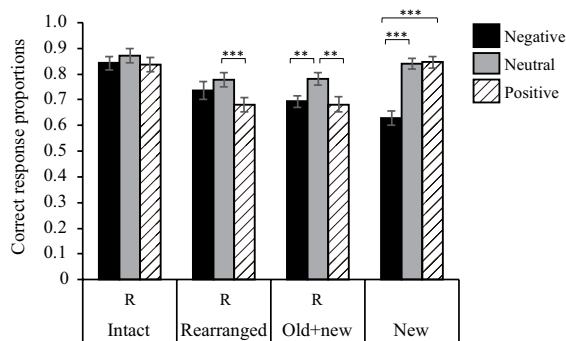
Regarding the factor of emotional valence in the two-way interaction, simple effect tests confirmed that the correct response proportions of intact pairs were insusceptible to this factor,  $ps \geq .262$ . In the rearranged condition, neutral pairs had much higher proportions than positive pairs ( $p < .001$ , 95% CI [0.051, 0.143]). In the old+new condition, neutral pairs had much higher proportions than both negative and positive ones ( $p = .002$ , 95% CI [0.029, 0.147];  $p = .001$ , 95% CI [0.040, 0.160]). Considering new cases, the proportions were much higher for both neutral and positive pairs than for negative ones ( $p < .001$ , 95% CI [0.160, 0.260];  $p < .001$ , 95% CI [0.162, 0.271]).

In sum, regarding the restudy condition, intact pairs had much higher correct response proportions than both rearranged and old+new pairs, which revealed that intact pairs were relatively easier to identify. Neutral pairs performed the best, then the positive pairs, while negative ones performed the worst, verifying a negative impairment effect. The correct response proportions behaved as the function of pair type by emotional valence.

## Sensitivity of Correct Response Proportions to the Learning Condition

The correct response proportions as the function of pair type by emotional valence by learning condition are displayed in Figure 4. The repeated-measures ANOVA of pair type by emotional valence by learning condition revealed a significant main effect of pair type,  $F(2, 58) = 55.438$ ,  $p < .001$ ,  $\eta_p^2 = .657$ ,  $\epsilon = 0.940$ , showing that intact pairs had much higher proportions than both rearranged and old+new pairs ( $p < .001$ , 95% CI [0.114, 0.202];  $p < .001$ , 95% CI [0.104, 0.177]). A reliable main effect of emotional valence was recorded,  $F(2, 58) = 25.173$ ,  $p < .001$ ,  $\eta_p^2 = .465$ ,  $\epsilon = 0.981$ , indicating the proportions were much higher for neutral pairs than for both negative and positive pairs ( $p < .001$ , 95% CI [0.041, 0.097];  $p < .001$ , 95% CI [0.035, 0.093]). The main



**FIGURE 3.**

Correct response proportions as the function of pair type by emotional valence in the restudy condition. R = restudy condition. Bars represent the SEM.

\*\*  $p < .01$ , \*\*\*  $p < .001$

effect of learning condition did not reach statistical significance,  $F(1, 29) = 0.508$ ,  $p = .482$ ,  $\eta_p^2 = .017$ ,  $\epsilon = 1.000$ .

However, pair type interacted with learning condition,  $F(2, 58) = 6.673$ ,  $p = .003$ ,  $\eta_p^2 = .187$ ,  $\epsilon = 0.927$ ; emotional valence interacted with learning condition,  $F(2, 58) = 7.672$ ,  $p = .001$ ,  $\eta_p^2 = .209$ ,  $\epsilon = 0.981$ ; the three-way interaction of pair type by emotional valence by learning condition was marginally significant,  $F(4, 116) = 2.308$ ,  $p = .069$ ,  $\eta_p^2 = .074$ ,  $\epsilon = 0.900$ . Simple effect tests for the three-way interaction demonstrated that, with regard to the learning condition, whether the intact pairs were negative or neutral, those experienced in the retrieval practice condition had much higher proportions than those experienced in the restudy condition ( $p = .048$ , 95% CI [0.000, 0.086];  $p = .044$ , 95% CI [0.001, 0.089]), demonstrating a reliable retrieval practice effect. Also, whether the rearranged pairs were negative or neutral, the proportions were much lower for pairs in the retrieval practice condition than in the restudy condition ( $p = .014$ , 95% CI [0.017, 0.133];  $p = .031$ , 95% CI [0.005, 0.101]). As to the old+new pairs, positive pairs under the retrieval practice condition yielded much higher proportions than those under the restudy condition ( $p = .023$ , 95% CI [0.009, 0.108]), illustrating a significant retrieval practice effect.

In sum, pairs under the retrieval practice condition did not always perform better than those under the restudy condition, indicating the retrieval practice effect was unstable and susceptible to both pair type and emotional valence.

## Analyses for Old Proportions and the Results

The above analyses showed that, under all three kinds of emotional valence, the correct response proportions of both rearranged and old+new pairs were much lower than those of intact pairs; in the positive case, the proportions of both rearranged and old+new pairs were much lower than those of new pairs. Thus, we examined whether participants tended to classify more rearranged and old+new pairs than new pairs as intact ones. To this end, we made a further analysis for old proportions (i.e., response proportions for different pairs as intact ones), similarly to

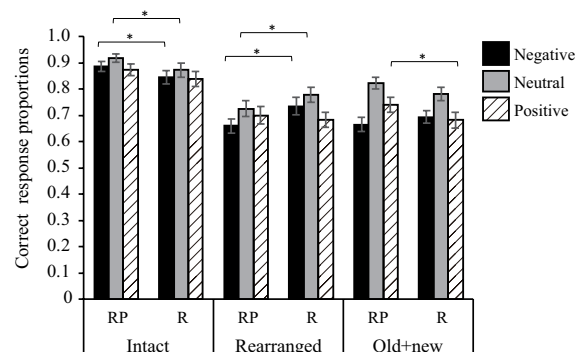
previous studies (Hockley et al., 2016; Nie & Jiang, 2019). The old proportions were submitted to a 4 (pair type: intact, rearranged, old+new, and new)  $\times$  3 (emotional valence: negative, neutral, and positive) repeated-measures ANOVA for retrieval practice and restudy conditions separately. Similarly, because we intended to test whether the old proportions behaved as a function of pair type, but considering that the current experiment was an incomplete within-subject design, it was essential to distinguish between the two learning conditions to ensure equal trial numbers in different pair types and thoroughly examine the impact of emotional valence on associative memory.

In addition, to examine whether the old proportions were sensitive to learning condition, they were submitted to a 3 (pair type: intact, rearranged, and old+new)  $\times$  2 (learning condition: retrieval practice and restudy)  $\times$  3 (emotional valence: negative, neutral, and positive) repeated-measures ANOVA.

## Analyses of Old Proportions in the Retrieval Practice Condition

Regarding the retrieval practice condition, the old proportions as the function of pair type by emotional valence are plotted in Figure 5. The repeated-measures ANOVA of pair type by emotional valence revealed a significant main effect of pair type,  $F(3, 87) = 1532.723$ ,  $p < .001$ ,  $\eta_p^2 = .981$ ,  $\epsilon = 0.488$ . Follow-up analyses for this main effect demonstrated much higher old proportions for intact pairs than the other three types of pairs ( $p < .001$ , 95% CI [0.744, 0.875]);  $p < .001$ , 95% CI [0.826, 0.929];  $p < .001$ , 95% CI [0.838, 0.937]), for rearranged pairs than both old+new and new pairs ( $p < .001$ , 95% CI [0.038, 0.098];  $p < .001$ , 95% CI [0.044, 0.113]), and for old+new pairs than new pairs ( $p = .042$ , 95% CI [0.000, 0.020]). The main effect of emotional valence did not reach statistical significance,  $F(2, 58) = 2.712$ ,  $p = .090$ ,  $\eta_p^2 = .086$ ,  $\epsilon = 0.776$ .

The two-way interaction reached statistical significance,  $F(6, 174) = 2.875$ ,  $p = .036$ ,  $\eta_p^2 = .090$ ,  $\epsilon = 0.542$ . Simple effect tests for the interaction revealed that, in the negative condition, much higher old proportions were found for intact pairs than the other three types of pairs ( $p < .001$ , 95% CI [0.715, 0.865];  $p < .001$ , 95% CI [0.802, 0.925];  $p < .001$ ,

**FIGURE 4.**

Correct response proportions as the function of pair type by emotional valence by learning condition. RP = retrieval practice; R = restudy condition. Bars represent the SEM.

\*  $p < .05$ .

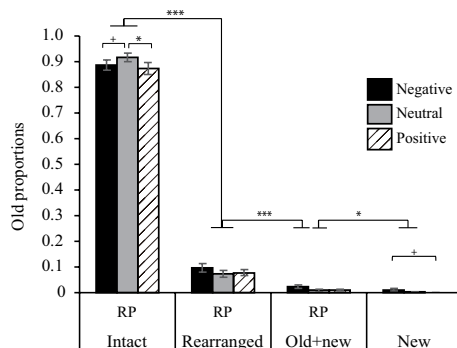
95% CI [0.819, 0.931]), and for rearranged pairs than both old+new and new pairs ( $p < .001$ , 95% CI [0.028, 0.118];  $p < .001$ , 95% CI [0.037, 0.133]). In the neutral condition, intact pairs produced much higher old proportions than the other three types of pairs ( $p < .001$ , 95% CI [0.777, 0.909];  $p < .001$ , 95% CI [0.862, 0.952];  $p < .001$ , 95% CI [0.869, 0.961]), and the old proportions were much higher for rearranged pairs than for both old+new and new pairs ( $p < .001$ , 95% CI [0.029, 0.097];  $p < .001$ , 95% CI [0.032, 0.111]). Regarding the positive condition, much higher old proportions were found for intact ones than for the other three types of pairs ( $p < .001$ , 95% CI [0.715, 0.875];  $p < .001$ , 95% CI [0.794, 0.933];  $p < .001$ , 95% CI [0.807, 0.939]), and for rearranged pairs than for both old+new and new pairs ( $p < .001$ , 95% CI [0.035, 0.102];  $p < .001$ , 95% CI [0.044, 0.113]).

Turning to the factor of emotional valence in the two-way interaction, simple effect tests indicated that, regarding the intact pairs, much higher old proportions were shown in neutral ones than in both negative and positive ones ( $p = .052$ , 95% CI [0.000, 0.060];  $p = .037$ , 95% CI [0.002, 0.085]). The old proportions of both rearranged and old+new pairs did not significantly differ across the three types of emotional valence,  $ps \geq .219$ . In new pairs, the old proportions were marginally higher for negative pairs than for positive ones ( $p = .051$ , 95% CI [0.000, 0.023]).

In sum, in the retrieval practice condition, rearranged pairs were the most likely to be classified as intact ones, followed by the old+new and new pairs, which all suggested that rearranged pairs could trigger the highest false alarm rate and the discriminations of such pairs were the hardest; furthermore, such patterns held true across the pairs of negative, neutral, and positive valences. The old proportions acted as the function of pair type by emotional valence, bringing about marginal significant negative impairment effect in new pairs.

## Analyses of Old Proportions in the Restudy Condition

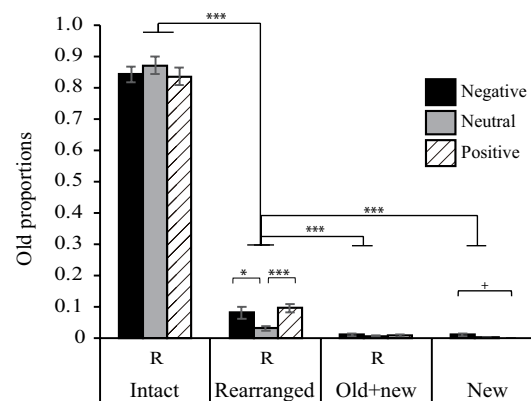
Regarding the restudy condition, the old proportions as the function of pair type by emotional valence are shown in Figure 6. The repeated-measures ANOVA of pair type by emotional valence confirmed a



**FIGURE 5.**

Old proportions as the function of pair type by emotional valence in the retrieval practice condition. RP = retrieval practice. Bars represent the SE.

\*  $p < .05$ , \*\*\*  $p < .001$ .



**FIGURE 6.**

Old proportions as the function of pair type by emotional valence in the restudy condition. R = restudy condition. Bars represent the SEM.

\*  $p < .05$ , \*\*\*  $p < .001$ .

significant main effect of pair type,  $F(3, 87) = 829.637$ ,  $p < .001$ ,  $\eta_p^2 = .966$ ,  $\epsilon = 0.410$ , which illustrated much higher old proportions for intact pairs than for the other three types of pairs ( $p < .001$ , 95% CI [0.695, 0.866];  $p < .001$ , 95% CI [0.768, 0.915];  $p < .001$ , 95% CI [0.774, 0.918]), and for rearranged pairs than for both old+new and new pairs ( $p < .001$ , 95% CI [0.034, 0.089];  $p < .001$ , 95% CI [0.035, 0.096]). The main effect of emotional valence was not statistically significant,  $F(2, 58) = 0.881$ ,  $p = .399$ ,  $\eta_p^2 = .029$ ,  $\epsilon = 0.791$ .

Further, pair type interacted with emotional valence,  $F(6, 174) = 5.976$ ,  $p = .001$ ,  $\eta_p^2 = .171$ ,  $\epsilon = 0.499$ . Simple effect tests for this interaction showed that, in the negative condition, intact pairs evoked much higher old proportions than the other three types of pairs ( $p < .001$ , 95% CI [0.672, 0.851];  $p < .001$ , 95% CI [0.757, 0.907];  $p < .001$ , 95% CI [0.756, 0.908]), with rearranged pairs having much higher old proportions than both old+new and new pairs ( $p = .003$ , 95% CI [0.019, 0.121];  $p = .003$ , 95% CI [0.019, 0.121]). In the neutral condition, intact pairs had much higher old proportions than the other three types of pairs ( $p < .001$ , 95% CI [0.751, 0.929];  $p < .001$ , 95% CI [0.783, 0.947];  $p < .001$ , 95% CI [0.789, 0.951]), and rearranged pairs had much higher old proportions than both old+new and new pairs ( $p = .014$ , 95% CI [0.004, 0.046];  $p = .009$ , 95% CI [0.006, 0.054]). In the positive condition, much higher old proportions were found for intact pairs than for the other three types of pairs ( $p < .001$ , 95% CI [0.637, 0.843];  $p < .001$ , 95% CI [0.743, 0.913];  $p < .001$ , 95% CI [0.754, 0.919]), with those for rearranged pairs being higher than for both old+new and new pairs ( $p < .001$ , 95% CI [0.053, 0.123];  $p < .001$ , 95% CI [0.058, 0.135]).

Considering the factor of emotional valence in the two-way interaction, none of the old proportions of intact and old+new pairs differed across the three types of emotional valence,  $ps \geq .262$ . In the rearranged condition, much higher old proportions were elicited in both negative and positive pairs than in neutral ones ( $p = .023$ , 95% CI [0.006, 0.094];  $p < .001$ , 95% CI [0.031, 0.099]). For new pairs, much higher old proportions were discovered for negative ones than for positive ones ( $p = .051$ , 95% CI [0.000, 0.023]).

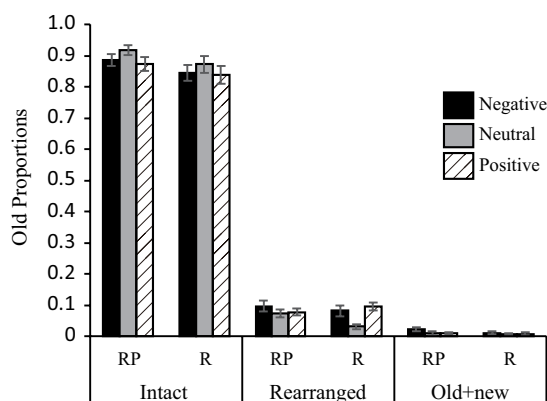
Thus, in the restudy condition, rearranged pairs were the most likely to be classified as intact ones, followed by the old+new and new pairs, which all demonstrated that the rearranged pairs tended to elicit the highest false alarm rate, and such patterns held true across pairs of negative, neutral, and positive valence. The old proportions behaved as the function of pair type by emotional valence, which resulted in a significant negative impairment effect in rearranged pairs and marginally significant negative impairment effect in new pairs.

## Sensitivity of Old Proportions to Learning Condition

Figure 7 shows the old proportions as the function of pair type by emotional valence by learning condition. The 3 (pair type: intact, rearranged, and old+new)  $\times$  2 (learning condition: retrieval practice and restudy)  $\times$  3 (emotional valence: negative, neutral, and positive) repeated-measures ANOVA found a reliable main effect of pair type,  $F(2, 58) = 1230.841, p < .001, \eta_p^2 = .977, \epsilon = 0.576$ , indicating the old proportions were much higher for intact pairs than both rearranged and old+new pairs ( $p < .001, 95\% \text{ CI } [0.732, 0.858]$ ;  $p < .001, 95\% \text{ CI } [0.807, 0.912]$ ), and also for rearranged pairs than old+new pairs ( $p < .001, 95\% \text{ CI } [0.043, 0.087]$ ). The main effect of learning condition was confirmed,  $F(1, 29) = 7.664, p = .010, \eta_p^2 = .209, \epsilon = 1.000$ , showing much higher old proportions in the retrieval practice condition than in the restudy condition ( $p = .010, 95\% \text{ CI } [0.005, 0.035]$ ). The main effect of emotional valence was not statistically significant,  $F(2, 58) = 0.620, p = .521, \eta_p^2 = .021, \epsilon = 0.871$ . Further, the two-way and three-way interactions were all not statistically significant,  $ps \geq .104$ .

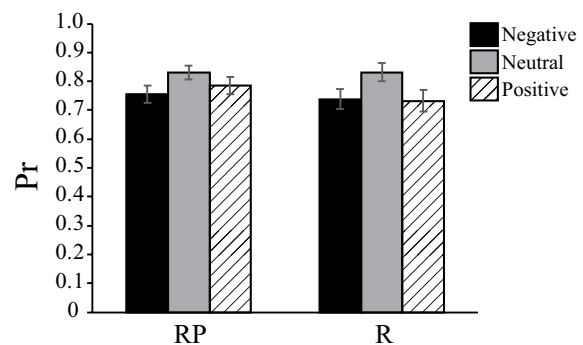
In brief, pairs under the retrieval practice condition might elicit much higher old proportions than those in the restudy condition.

As the above results showed that participants were inclined to identify all the rearranged, old+new, and new pairs as intact ones, we further analyzed the discrimination index of Pr (i.e., the hit proportions of intact pairs minus the false alarm proportions of the other three types of pairs, see Figure 8 for each condition), as prior literature on associative memory did (Bridger et al., 2017; Nie & Jiang, 2019). To examine



**FIGURE 7.**

Old proportions as the function of pair type by emotional valence by learning condition. RP = retrieval practice condition; R = restudy condition. Bars represent the SEM.



**FIGURE 8.**

Discrimination index of Pr as the function of learning condition by emotional valence. RP = retrieval practice condition; R = restudy condition. Bars represent the SEM.

whether the Pr was susceptible to our manipulated variables, a 3 (emotional valence: negative, neutral, and positive)  $\times$  2 (learning condition: retrieval practice, and restudy) repeated-measures ANOVA was conducted. It revealed a significant main effect of emotional valence,  $F(2, 58) = 9.783, p < .001, \eta_p^2 = .252, \epsilon = 0.987$ , illustrating that neutral pairs produced a much higher Pr than both negative and positive ones ( $p = .002, 95\% \text{ CI } [0.030, 0.140]$ ;  $p = .004, 95\% \text{ CI } [0.020, 0.126]$ ). The main effect of learning condition was not statistically significant,  $F(1, 29) = 1.249, p = .273, \eta_p^2 = .041, \epsilon = 1.000$ . These findings demonstrated that Pr was merely modulated by emotional valence.

## Analyses for Response Times and the Data

To detect whether there was a trade-off between correct response proportions and RTs, we conducted the same repeated-measures ANOVA of pair type by emotional valence for RTs in the retrieval practice and restudy conditions separately, similarly to the correct response proportions. Also, because we intended to explore whether the RTs behaved as a function of pair type, while the current experiment was an incomplete within-subject design, it was necessary to draw a distinction between the two learning conditions to keep a balance among the trial number of pair type and comprehensively investigate the effect of emotional valence on associative memory. Moreover, to test whether the RTs were sensitive to the factor of learning condition, a repeated-measures ANOVA of pair type by emotional valence by learning condition was performed.

## Analyses of the Response Times in Retrieval Practice Condition

The RTs for the correctly identified pairs, as the function of pair type by emotional valence in retrieval practice condition, are displayed in Figure 9. Under this condition, the 4 (pair type: intact, rearranged, old+new, and new)  $\times$  3 (emotional valence: negative, neutral, and positive) repeated-measures ANOVA confirmed a significant main effect of pair type,  $F(3, 87) = 131.712, p < .001, \eta_p^2 = .820, \epsilon = 0.781$ , which showed much shorter RTs for intact pairs than the other three types of pairs ( $p < .001, 95\% \text{ CI } [-695.051, -502.001]$ ;  $p < .001, 95\% \text{ CI } [-649.652,$

−493.087];  $p < .001$ , 95% CI [−594.316, −392.791]). The main effect of emotional valence reached statistical significance,  $F(2, 58) = 38.865$ ,  $p < .001$ ,  $\eta_p^2 = .573$ ,  $\epsilon = 0.885$ , revealing faster RTs for neutral pairs than for both negative and positive ones ( $p < .001$ , 95% CI [−198.472, −105.170];  $p = .014$ , 95% CI [−80.555, −7.544]), and also for positive pairs than negative ones ( $p < .001$ , 95% CI [−158.484, −57.059]).

Additionally, pair type interacted with emotional valence,  $F(6, 174) = 8.818$ ,  $p < .001$ ,  $\eta_p^2 = .233$ ,  $\epsilon = 0.658$ . Simple effect tests for the interaction showed that in the negative condition, participants responded much faster to intact pairs than to the other three types of pairs ( $p < .001$ , 95% CI [−706.235, −470.375];  $p < .001$ , 95% CI [−697.357, −485.526];  $p < .001$ , 95% CI [−784.185, −491.948]). In the neutral condition, intact pairs responded faster than the other three types of pairs ( $p < .001$ , 95% CI [−714.240, −517.754];  $p < .001$ , 95% CI [−689.878, −500.427];  $p < .001$ , 95% CI [−574.770, −341.606]), and new pairs than both rearranged and old+new ones ( $p = .038$ , 95% CI [−309.436, −6.182];  $p = .001$ , 95% CI [−224.036, −49.894]). In the positive condition, the RTs were much shorter for intact pairs than for the other three types of pairs ( $p < .001$ , 95% CI [−720.863, −461.690];  $p < .001$ , 95% CI [−615.901, −439.127];  $p < .001$ , 95% CI [−495.747, −273.068]), and were much shorter for new pairs than for both rearranged and old+new pairs ( $p = .002$ , 95% CI [−347.646, −66.093];  $p = .001$ , 95% CI [−238.265, −47.948]).

Regarding the factor of emotional valence in the two-way interaction, simple effect tests showed that, with respect to the intact pairs, participants responded much faster to neutral ones than to both negative and positive ones ( $p < .001$ , 95% CI [−149.491, −79.913];  $p = .004$ , 95% CI [−146.233, −24.935]). For rearranged pairs, neutral ones had much shorter RTs than negative ones ( $p = .035$ , 95% CI [−168.980, −5.040]). For the old+new case, both neutral and positive pairs had much faster RTs than negative ones ( $p = .006$ , 95% CI [−194.695, −27.288];  $p = .040$ , 95% CI [−182.784, −3.308]). Lastly, both neutral and positive new pairs had much faster RTs than negative new ones ( $p < .001$ , 95% CI [−409.923, −179.238];  $p < .001$ , 95% CI [−404.350, −161.204]).

To sum up, in the retrieval practice condition, the intact pairs were the fastest ones to be identified, indicating that the identifications of in-

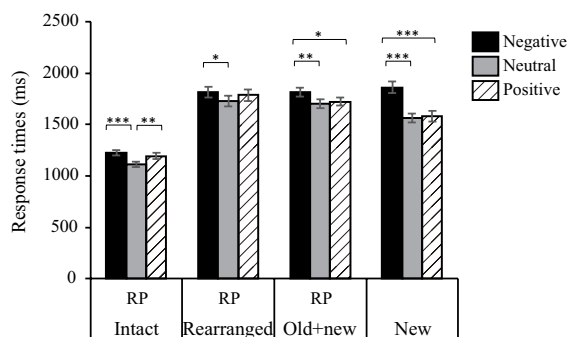
tact pairs were the easiest. Neutral pairs tended to have faster RTs than positive pairs, while negative pairs had the slowest RTs, confirming a negative impairment effect. The RTs could be affected by the interaction of pair type by emotional valence.

## Analyses of the Response Times in the Restudy Condition

For the restudy condition, the RTs as the function of pair type by emotional valence are presented in Figure 10. The 4 (pair type: intact, rearranged, old+new, and new)  $\times$  3 (emotional valence: negative, neutral, and positive) repeated-measures ANOVA for RTs detected a significant main effect of pair type,  $F(3, 87) = 171.543$ ,  $p < .001$ ,  $\eta_p^2 = .855$ ,  $\epsilon = 0.713$ , demonstrating that intact pairs had much faster RTs than the other three types of pairs ( $p < .001$ , 95% CI [−695.826, −530.224];  $p < .001$ , 95% CI [−724.917, −570.584];  $p < .001$ , 95% CI [−438.011, −238.392]), and that new pairs had faster RTs than both rearranged and old+new pairs ( $p < .001$ , 95% CI [−398.797, −150.850];  $p < .001$ , 95% CI [−393.528, −225.570]). A significant main effect of emotional valence was found,  $F(2, 58) = 27.601$ ,  $p < .001$ ,  $\eta_p^2 = .488$ ,  $\epsilon = 0.892$ , demonstrating that neutral pairs had much faster RTs than negative and positive ones ( $p < .001$ , 95% CI [−171.974, −81.706];  $p < .001$ , 95% CI [−96.909, −26.072]), and positive pairs had faster RTs than negative ones ( $p = .006$ , 95% CI [−113.877, −16.820]).

Pair type interacted with emotional valence,  $F(6, 174) = 13.129$ ,  $p < .001$ ,  $\eta_p^2 = .312$ ,  $\epsilon = 0.742$ . Simple effect tests revealed that in the negative case, intact pairs had much faster RTs than the other three types of pairs ( $p < .001$ , 95% CI [−630.930, −382.306];  $p < .001$ , 95% CI [−722.394, −476.458];  $p < .001$ , 95% CI [−600.159, −315.462]), and rearranged pairs had much shorter RTs compared to old+new pairs ( $p = .025$ , 95% CI [−177.380, −8.235]). In the neutral condition, participants responded much faster to intact pairs than to the other three types of pairs ( $p < .001$ , 95% CI [−679.462, −489.471];  $p < .001$ , 95% CI [−716.578, −551.433];  $p < .001$ , 95% CI [−379.848, −163.416]), and to new pairs than to both rearranged and old+new pairs ( $p < .001$ , 95% CI [−454.130, −171.538];  $p < .001$ , 95% CI [−455.118, −269.630]). In the positive condition, intact pairs had much faster RTs than the other three types of pairs ( $p < .001$ , 95% CI [−865.340, −630.642];  $p < .001$ , 95% CI [−824.531, −595.109];  $p < .001$ , 95% CI [−418.766, −151.557]), and new pairs had much faster RTs than both rearranged and old+new pairs ( $p < .001$ , 95% CI [−615.324, −310.334];  $p < .001$ , 95% CI [−534.845, −314.472]).

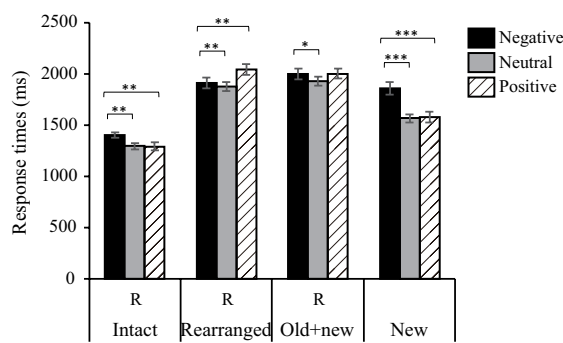
Regarding the factor of emotional valence in the two-way interaction, simple effect tests indicated that, in the intact condition, both neutral and positive pairs had much faster RTs than negative ones ( $p = .001$ , 95% CI [−174.750, −42.055];  $p = .003$ , 95% CI [−187.276, −32.980]). In the rearranged condition, both negative and neutral ones had much faster RTs than positive ones ( $p = .003$ , 95% CI [−221.074, −41.415];  $p = .001$ , 95% CI [−257.176, −66.421]). In the old+new condition, neutral pairs had much faster RTs than negative ones ( $p = .041$ , 95% CI [−145.288, −2.356]). In the new condition, both neutral and positive pairs received much faster RTs than negative ones ( $p < .001$ , 95% CI [−409.923, −179.238];  $p < .001$ , 95% CI [−404.350, −161.204]).



**FIGURE 9.**

Response times as the function of pair type by emotional valence in retrieval practice condition. RP = retrieval practice condition. Bars represent the SEM.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

**FIGURE 10.**

Response times as the function of pair type by emotional valence in the restudy condition. R = restudy condition. Bars represent the SEM.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

In brief, in the restudy condition, the identifications of intact pairs were the fastest, then the new pairs, with rearranged and old+new pairs having the slowest RTs. Such patterns illustrated that the identifications of intact and new pairs were far easier than the others. Neutral pairs tended to have much faster RTs than positive pairs, which were followed by negative pairs, illuminating a negative impairment effect. The RTs could be modulated by the interaction of pair type by emotional valence.

### Sensitivity of Response Times to Learning Condition

The response times as the function of pair type by emotional valence by learning condition are plotted in Figure 11. The 3 (pair type: intact, rearranged, and old+new)  $\times$  2 (learning condition: retrieval practice and restudy)  $\times$  3 (emotional valence: negative, neutral, and positive) repeated-measures ANOVA confirmed a significant main effect of pair type,  $F(2, 58) = 414.903$ ,  $p < .001$ ,  $\eta_p^2 = .935$ ,  $\epsilon = 0.978$ , which indicated that intact pairs had much faster RTs than both rearranged and old+new pairs ( $p < .001$ , 95% CI  $[-671.344, -540.207]$ ;  $p < .001$ , 95% CI  $[-667.081, -552.039]$ ). There was a significant main effect of learning condition,  $F(1, 29) = 63.692$ ,  $p < .001$ ,  $\eta_p^2 = .687$ ,  $\epsilon = 1.000$ , showing that the pairs in retrieval practice condition had much faster RTs than those in the restudy condition ( $p < .001$ , 95% CI  $[-233.221, -138.070]$ ), which demonstrated a reliable retrieval practice effect. The main effect of emotional valence was significant as well,  $F(2, 58) = 18.498$ ,  $p < .001$ ,  $\eta_p^2 = .389$ ,  $\epsilon = 0.849$ , indicating that neutral pairs had much faster RTs than both negative and positive pairs ( $p < .001$ , 95% CI  $[-127.041, -48.119]$ ;  $p < .001$ , 95% CI  $[-96.129, -36.723]$ ).

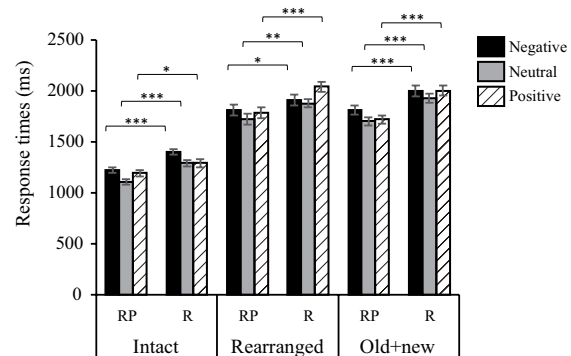
In addition, the interaction between pair type and learning condition was marginally significant,  $F(2, 58) = 2.940$ ,  $p = .064$ ,  $\eta_p^2 = .092$ ,  $\epsilon = 0.951$ ; the interaction between emotional valence and learning condition was significant,  $F(2, 58) = 3.390$ ,  $p = .045$ ,  $\eta_p^2 = .105$ ,  $\epsilon = 0.912$ ; the three-way interaction of pair type by emotional valence by learning condition also reached statistical significance,  $F(4, 116) = 4.462$ ,  $p = .004$ ,  $\eta_p^2 = .133$ ,  $\epsilon = 0.824$ . Simple effect tests for the three-way interaction demonstrated that, with regard to the learning condi-

tion, regardless of the emotional valences of intact pairs, the ones in the retrieval practice condition had much faster RTs than those in the restudy condition ( $p < .001$ , 95% CI  $[-239.313, -121.200]$ ;  $p < .001$ , 95% CI  $[-236.798, -136.313]$ ;  $p = .017$ , 95% CI  $[-179.524, -18.967]$ ), and rearranged pairs ( $p = .032$ , 95% CI  $[-187.927, -9.211]$ ;  $p = .001$ , 95% CI  $[-238.705, -71.345]$ ;  $p < .001$ , 95% CI  $[-362.316, -149.604]$ ) as well as old+new pairs ( $p < .001$ , 95% CI  $[-270.326, -106.154]$ ;  $p < .001$ , 95% CI  $[-290.827, -159.990]$ ;  $p < .001$ , 95% CI  $[-362.724, -200.380]$ ) shared the same pattern as intact ones.

In sum, the pairs in retrieval practice condition always had much faster RTs than those in the restudy condition, indicating a stable retrieval practice effect from the perspective of RTs, and the effect was not be modulated by pair type or emotional valence.

### DISCUSSION

The current study evaluated how emotional valence affects associative memory from the perspective of retrieval practice. Most of our expectations were verified. First, the performance of intact pairs was the best in both the retrieval practice and restudy conditions, and rearranged pairs were the most likely to be classified as intact ones in both conditions, demonstrating that the involvement of recollection-driven process varied with pair type. The responses were the fastest for intact pairs, then for new pairs, while rearranged and old+new pairs were the slowest. Second, neutral pairs had much shorter response times than positive pairs, which were followed by negative pairs, demonstrating a negative impairment effect; this was insensitive to the learning condition, and this negative impairment effect corroborated the account of spontaneous interactive imagery. Also, the Pr was affected by the emotional valence of stimuli but not by the learning condition. Lastly, there was a significant retrieval practice effect, which was susceptible to both emotional valence and pair type. The implications of the findings and possible accounts concerning these findings for the main loci of interest are discussed. Also, implications for future research directions are offered.

**FIGURE 11.**

Response times as the function of pair type by emotional valence by learning condition. RP = retrieval practice condition; R = restudy condition. Bars represent the SEM.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .



## Involvement of Recollection-Driven Process Varies with Pair Type

As outlined in the Introduction, the associative memory relied differently upon the familiarity-based process and the recollection-based process. Some researchers suggested that intact and rearranged pairs are engaged by familiarity-based processes to a similar extent, as the items of both pairs all came from previously studied pairs, making the differentiations of intact and rearranged pairs primarily dependent upon the recollection-based process. However, the discernments of old+new and new pairs were less dependent on the recollection-based process, and could be performed on the basis of merely familiarity-based process (Buchler et al., 2008, 2011; Nie & Jiang, 2019). In terms of correct response proportions, the current intact pairs performed better than rearranged ones in both the retrieval practice and restudy conditions, which not only confirms that the involvement of the recollection-based process differs between these two types of pairs, but also demonstrates that the actively conscious process (i.e., effort) elicited by retrieval practice or the re-exposure to studied pairs in the restudy condition may enhance the availability and the accessibility of intact pairs. In addition, the benefit of intact pairs could also be authenticated by RTs, as the intact pairs could be responded to much faster than the other three types of pairs, and such a manner held true in both retrieval practice and restudy conditions, demonstrating that the correct responses and RTs exhibit a tendency of trade-off. Considering associative memory, these findings provide compelling evidence for dual-process models (Bridger et al., 2017; Li et al., 2019; Liu et al., 2021; Maylor et al., 2019; Nie et al., 2019; Nie & Jiang, 2019; Ye et al., 2019).

However, inconsistent with the assumption of much higher correct response proportions for old+new and new pairs than rearranged pairs, we confirmed similar correct proportions for rearranged, old+new, and new pairs. Thus, such findings did not dovetail with the tendency reported by previous studies (Buchler et al., 2008, 2011; Nie & Jiang, 2019), which found that intact and new pairs performed the best, then the old+new ones, while rearranged pairs performed the worst. One possible account for the discrepancy between the current study and the previous studies was the arrangement of the learning condition. In this condition, re-exposure to the studied pairs might make the resources of familiarity- and recollection-based processes unequal for the items of all pair types, resulting in similar engagements of familiarity- and recollection-based processes across rearranged, old+new, and new pairs, eventually producing similar availability and accessibility for these pair types.

However, when taking the old proportions and RTs into consideration, this perspective, that is, similar engagements of familiarity- and recollection-based processes across rearranged, old+new, and new pairs, would be unreasonable. Regarding old proportions, the analyses demonstrated that, compared with old+new and new ones, rearranged pairs were more likely to be classified as intact ones, and such a pattern held regardless of whether the items of rearranged pairs were previously under the retrieval practice condition or the restudy condition. These findings indicate that the rearranged pairs elicited the recollection-based process to a greater degree than old+new and new

ones, but that it was not strong enough to support participants to recall the paired items, consequently, leading to more false alarms for the former. Interestingly, these findings were fairly in line with the patterns reported in one of our previous studies (Nie & Jiang, 2019). Compared with new pairs, the old+new pairs in the retrieval practice condition were more likely to be identified as intact, showing that the memory traces of studied words in the old+new pairs tended to trigger the recollection-based process more, similarly, eliciting more false alarms for them. Turning to the RTs, rearranged and old+new pairs had much slower RTs than new pairs, but only in the restudy condition, which substantiated that the involvement of the familiarity- and recollection-based processes might differ between these pair types, similar to the previous investigations (Buchler et al., 2008, 2011; Nie & Jiang, 2019).

## Emotional Valence Plays a Vital Role in Associative Memory

As expected, the current study found that the performance of associative memory was sensitive to the factor of emotional valence of stimuli, which was consistent with previous findings (Caplan et al., 2019; Madan et al., 2017, 2019; Nadarevic, 2017; Nie & Jiang, 2019; Siegel et al., 2019; Zimmerman & Kelley, 2010). In terms of correct response proportions, both negative and positive valence were detrimental to associative memory, as neutral pairs performed better than emotional pairs (including both positive and negative ones), and positive pairs performed better than negative ones, demonstrating a negative impairment effect. The above findings are relatively congruent with those reported in previous research (Caplan et al., 2019; Madan et al., 2017; Nie & Jiang, 2019), and further corroborate the account of spontaneous interactive imagery (Caplan et al., 2019; Madan et al., 2017). According to this account, the inferiority of the negative condition can be attributed to the lower effectiveness of the spontaneous interactive imagery in the negative condition, while neutral stimuli can gain more unitization from the spontaneous interactive imagery.

Turning to old proportions, the negative impairment effect also occurred, and it was observed in intact and new pairs under the retrieval practice condition, and in both rearranged and new pairs under the restudy condition. This pattern confirmed that the intact and rearranged pairs whose retrieval relied more on recollection-based process and new pairs whose retrieval depended more on familiarity-based process tended to evoke the negative impairment effect. Also, as for the discrimination index of  $Pr$ , a negative impairment effect did occur. Together, these findings dovetailed with our previous research (Nie & Jiang, 2019), which suggested a vital role of negative valence in the attenuation of associative memory. In terms of RTs, faster responses for neutral pairs than emotional ones were revealed, regardless of condition, retrieval practice or restudy. Also, there was a trade-off between RTs and correct response proportions for the negative impairment effect. The above patterns indicate that this effect is stable.

Furthermore, when compared with previous studies, following conclusions can be drawn from the current study. First, the occurrence of the negative impairment effect was insusceptible to the pattern of the final test, as the current experiment utilized the associative recognition

task, while Madan et al. (2017) and Caplan et al. (2019) applied the manner of selective recognition. Second, the emergence of negative impairment was insensitive to the learning condition, mainly due to the fact that it was observed in both the retrieval practice and restudy conditions and the aforementioned trade-off held true under both conditions. Lastly, the presence of the positive bias was dependent, as it acted differently when the final test varied: It was absent in the current recognition case but was observed in the cued recall circumstance in studies by Madan et al. (2019) and Zimmerman and Kelley (2010), demonstrating that the broaden-and-build theory is inapplicable to the current study.

## The Emergence of Retrieval Practice Effect is Sensitive to the Interaction of Emotional Valence by Pair Type

With regard to the correct response proportions, a reliable retrieval practice effect was verified to some degree, which was consistent with previous findings (Abel & Roediger, 2017; Binks, 2018; Buchin & Mulligan, 2017; Cho, Neely, Brennan et al., 2017; Cho, Neely, Crocco et al., 2017; Peng et al., 2019; Peterson & Wissman, 2018; Rickard & Pan, 2018; Szöllösi et al., 2017), such that retrieval practice rather than restudy is an effective way to enhance memory performance. Further, pairs in the retrieval practice condition always received faster responses than those in the restudy condition, indicating a stable retrieval practice effect from the angle of RTs. Together, these findings support the desirable difficulty framework, which claims that the retrieval practice condition demands more effort than the restudy condition, resulting in the enhancement of the availability and accessibility for the items from the retrieval practice condition in the final test (Binks, 2018; Greving & Richter, 2018; Kubik et al., 2018; Rickard & Pan, 2018).

In contrast to previous studies (Emmerdinger et al., 2018; Emmerdinger & Kuhbandner, 2018; Jia et al., 2018) that did not reveal any modulation of emotional valence of stimuli on the retrieval practice effect, we revealed a significant sensitivity of this effect to the emotional valence of stimuli, but the contributions of emotional valence interacted with those of pair type: a significant effect was observed for intact pairs of negative and neutral valences, and also for old+new pairs of positive valence. For rearranged pairs, an opposite pattern was recorded, showing the correct response proportions were much lower in the retrieval practice condition than in the restudy condition. Thus, the contribution of emotional valence in retrieval practice effect is complicated. We have not yet found a plausible explanation for these differences, but have verified the susceptibility of the retrieval practice effect to emotional valence of stimuli.

Compared to previous research (Emmerdinger et al., 2018; Emmerdinger & Kuhbandner, 2018; Jia et al., 2018), the current sensitivity of the retrieval practice effect to emotional valence was due to the fact that our study met the following requirements: shorter delay interval, manipulation of feedback, and pure emotionality in our pairs. To be specific, the current study adopted a delay interval of 1 min, while the intervals in previous studies were much longer, such as twenty minutes, one week, two weeks, or even thirteen weeks (Emmerdinger et al.,

2018; Emmerdinger & Kuhbandner, 2018; Jia et al., 2018). Next, we set up feedback in the retrieval practice condition, in contrast to no feedback in previous research (Emmerdinger et al., 2018; Emmerdinger & Kuhbandner, 2018; Jia et al., 2018). Lastly, pairs of pure emotionality were controlled in the current study, while previous investigations applied mixed emotional ones (Emmerdinger et al., 2018; Jia et al., 2018). We must admit that it is not yet possible to distinguish which manipulation plays the most important role in eliciting the sensitivity of the retrieval practice effect to emotional valence. Thus, further studies are needed.

## CONCLUSION

In conclusion, the current study substantiates that the involvement of the recollection-driven process varies with pair type, providing telling evidence for dual-process models. A negative impairment effect was observed, and its occurrence was regardless of pair type and learning condition, which confirms the account of spontaneous interactive imagery. Moreover, the benefit of retrieval practice is dependent upon the interaction of emotional valence and pair type.

## ACKNOWLEDGEMENTS

The authors declare that they have no conflict of interest.

All authors have seen and approved the final version of the manuscript being submitted. We warrant that the article is our original work, it has not received prior publication and is not under consideration for publication elsewhere.

This study is supported by the Projects of Zhejiang Provincial Natural Science Foundation of China (Grant No. LY21C090002), of Zhejiang Federation of Humanities and Social Sciences Circles (Grant No. 2021N78), of Humanities and Social Sciences, Ministry of Education of China (Grant No. 17YJA190010), of National Natural Science Foundation of China (Grant No. 31300831), and the Fundamental Research Funds for the Central Universities.

## REFERENCES

- Abel, M., & Roediger, H. L. (2017). Comparing the testing effect under blocked and mixed practice: The mnemonic benefits of retrieval practice are not affected by practice format. *Memory & Cognition*, 45, 81–92. doi: 10.3758/s13421-016-0641-8
- Ahmad, F. N., & Hockley, W. E. (2014). The role of familiarity in associative recognition of unitized compound word pairs. *Quarterly Journal of Experimental Psychology*, 67, 2301–2324. doi: 10.1080/17470218.2014.923007
- Baadte, C., & Meinhardt-Injac, B. (2019). The picture superiority effect in associative memory: A developmental study. *British Journal of Developmental Psychology*, 37, 382–395. doi: 10.1111/bjdp.12280
- Bae, C. L., Theriault, D. J., & Redifer, J. L. (2019). Investigating the testing effect: Retrieval as a characteristic of effective study strategies. *Learning and Instruction*, 60, 206–214. doi: 10.1016/j.learninstruc.2017.12.008
- Beardsley, M., Hernández-Leo, D., & Ramirez-Melendez, R. (2018).

- Seeking reproducibility: Assessing a multimodal study of the testing effect. *Journal of Computer Assisted Learning*, 34, 378–386. doi: 10.1111/jcal.12265
- Bellander, M., Eschen, A., Lövdén, M., Martin, M., Bäckman, L., & Brehmer, Y. (2017). No evidence for improved associative memory performance following process-based associative memory training in older adults. *Frontiers in Aging Neuroscience*, 8, 326. doi: 10.3389/fnagi.2016.00326
- Berger, N., Crossman, M., & Brandt, K. R. (2018). No evidence for age-related differences in item-method directed forgetting of emotional words. *Quarterly Journal of Experimental Psychology*, 71, 595–604. doi: 10.1080/17470218.2016.1264433
- Binks, S. (2018). Testing enhances learning: A review of the literature. *Journal of Professional Nursing*, 34, 205–210. doi: 10.1016/j.profnurs.2017.08.008
- Bisby, J. A., & Burgess, N. (2014). Negative affect impairs associative memory but not item memory. *Learning & Memory*, 21, 21–27. doi: 10.1101/lm.032409.113
- Bradley, M. M., & Lang, P. J. (1999). Affective Norms for English Words (ANEW): Instruction manual and affective ratings. *Journal of Royal Microscopica*, 88, 630–634. doi: 10.1111/j.1365-2621.2000.tb16063.xx
- Bridger, E. K., Kursawe, A., Bader, R., Tibon, R., Gronau, N., Levy, D. A., & Mecklinger, A. (2017). Age effects on associative memory for novel picture pairings. *Brain Research*, 1664, 102–115. doi: 10.1016/j.brainres.2017.03.031
- Buchin, Z. L., & Mulligan, N. W. (2017). The testing effect under divided attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 1934–1947. doi: 10.1037/xlm0000427
- Buchler, N. G., Faunce, P., Light, L. L., Gottfredson, N., & Reder, L. M. (2011). Effects of repetition on associative recognition in young and older adults: Item and associative strengthening. *Psychology and Aging*, 26, 111–126. doi: 10.1037/a0020816
- Buchler, N. G., Light, L. L., & Reder, L. M. (2008). Memory for items and associations: Distinct representations and processes in associative recognition. *Journal of Memory and Language*, 59, 183–199. doi: 10.1016/j.jml.2008.04.001
- Caplan, J. B., Sommer, T., Madan, C. R., & Fujiwara, E. (2019). Reduced associative memory for negative information: Impact of confidence and interactive imagery during study. *Cognition and Emotion*, 33, 1745–1753. doi: 10.1080/02699931.2019.1602028
- Cho, K. W., Neely, J. H., Brennan, M. K., Vitrano, D., & Crocco, S. (2017). Does testing increase spontaneous mediation in learning semantically related paired associates? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 1768–1778. doi: 10.1037/xlm0000414
- Cho, K. W., Neely, J. H., Crocco, S., & Vitrano, D. (2017). Testing enhances both encoding and retrieval for both tested and untested items. *Quarterly Journal of Experimental Psychology*, 70, 1211–1235. doi: 10.1080/17470218.2016.1175485
- De Brigard, F., Langella, S., Stanley, M. L., Castel, A. D., & Giovanello, K. S. (2020). Age-related differences in recognition in associative memory. *Aging, Neuropsychology, and Cognition*, 27, 289–301. doi: 10.1080/13825585.2019.1607820
- Delhay, E., Folville, A., & Bastin, C. (2019). The impact of semantic relatedness on associative memory in aging depending on the semantic relationships between the memoranda. *Experimental Aging Research*, 45, 469–479. doi: 10.1080/0361073X.2019.1664442
- Emmerdinger, K. J., Kuhbandner, C., & Berchtold, F. (2018). Testing emotional memories: Does negative emotional significance influence the benefit received from testing? *Cognition and Emotion*, 32, 852–859. doi: 10.1080/02699931.2017.1359496
- Emmerdinger, K. J., & Kuhbandner, C. (2018). Testing memories of personally experienced events: The testing effect seems not to persist in autobiographical memory. *Frontiers in Psychology*, 9, 810. doi: 10.3389/fpsyg.2018.00810
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. doi: 10.3758/BRM.41.4.1149
- Fredrickson, B. L. (2013). Positive emotions broaden and build. *Advances in Experimental Social Psychology*, 47, 1–53. doi: 10.1016/B978-0-12-407236-7.00001-2
- Greving, S., & Richter, T. (2018). Examining the testing effect in university teaching: Retrieval and question format matter. *Frontiers in Psychology*, 9, 2412. doi: 10.3389/fpsyg.2018.02412
- Hockley, W. E., Ahmad, F. N., & Nicholson, R. (2016). Intentional and incidental encoding of item and associative information in the directed forgetting procedure. *Memory & Cognition*, 44, 220–228. doi: 10.3758/s13421-015-0557-8
- Huff, M. J., Yates, T. J., & Balota, D. A. (2018). Evaluating the contributions of task expectancy in the testing and guessing benefits on recognition memory. *Memory*, 26, 1065–1083. doi: 10.1080/09658211.2018.1467929
- Huguet, M., Payne, J. D., Kim, S. Y., & Alger, S. E. (2019). Overnight sleep benefits both neutral and negative direct associative and relational memory. *Cognitive, Affective, & Behavioral Neuroscience*, 19, 1391–1403. doi: 10.3758/s13415-019-00746-8
- Jia, X., Gao, C., Cui, L., & Guo, C. (2018). Does emotion arousal influence the benefit received from testing: Insights from neural correlates of retrieval mode effect. *NeuroReport*, 29, 1449–1455. doi: 10.1097/WNR.0000000000001130
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126, 349–370.
- Kubik, V., Jönsson, F. U., Knopf, M., & Mack, W. (2018). The direct testing effect is pervasive in action memory: Analyses of recall accuracy and recall speed. *Frontiers in Psychology*, 9, 1632. doi: 10.3389/fpsyg.2018.01632
- Leach, R. C., McCurdy, M. P., Trumbo, M. C., Matzen, L. E., & Leshikar, E. D. (2019). Differential age effects of transcranial direct current stimulation on associative memory. *Journals of Gerontology: Psychological Sciences*, 74, 1163–1173. doi: 10.1093/geronb/gby003
- Li, B., Han, M., Guo, C., & Tibon, R. (2019). Unitization modulates recognition of within-domain and cross-domain associations:

- Evidence from event-related potentials. *Psychophysiology*, 56, online. doi: 10.1111/psyp.13446
- Liu, Y., Rosburg, T., Gao, C., Weber, C., & Guo, C. (2017). Differentiation of subsequent memory effects between retrieval practice and elaborative study. *Biological Psychology*, 127, 134–147. doi: 10.1016/j.biopsycho.2017.05.010
- Liu, Z., Wu, J., & Guo, C. (2021). The effect of unitization on associative recognition was not moderated by the unitization-congruence between original and rearranged picture pairs (UC) for picture stimuli. *Psychological Research*, 85, 268–279. doi: 10.1007/s00426-019-01260-9
- Lucas, H. D., Gupta, R. S., Hubbard, R. J., & Federmeier, K. D. (2019). Adult age differences in the use of conceptual combination as an associative encoding strategy. *Frontiers in Human Neuroscience*, 13, 339. doi: 10.3389/fnhum.2019.00339
- Madan, C. R., Fujiwara, E., Caplan, J. B., & Sommer, T. (2017). Emotional arousal impairs association-memory: Roles of amygdala and hippocampus. *NeuroImage*, 156, 14–28. doi: 10.1016/j.neuroimage.2017.04.065
- Madan, C. R., Scott, S. M. E., & Kensinger, E. A. (2019). Positive emotion enhances association-memory. *Emotion*, 19, 733–740. doi: 10.1037/emo0000465
- Maylor, E. A., Long, H. R., & Newstead, R. A. (2019). Differential effects of alcohol on associative versus item memory. *Applied Cognitive Psychology*, 33, 386–392. doi: 10.1002/acp.3476
- Minear, M., Coane, J. H., Boland, S. C., Cooney, L. H., & Albat, M. (2018). The benefits of retrieval practice depend on item difficulty and intelligence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44, 1474–1486. doi: 10.1037/xlm0000486
- Mulligan, N. W., Susser, J. A., & Smith, S. A. (2016). The testing effect is moderated by experimental design. *Journal of Memory and Language*, 90, 49–65. doi: 10.1016/j.jml.2016.03.006
- Nadarevic, L. (2017). Emotionally enhanced memory for negatively arousing words: Storage or retrieval advantage? *Cognition and Emotion*, 31, 1557–1570. doi: 10.1080/02699931.2016.1242477
- Nie, A., & Jiang, G. (2019). Does stimulus emotionality influence associative memory? Insights from directed forgetting. *Current Psychology*, 2019. doi: 10.1007/s12144-019-00449-w
- Nie, A., Ke, C., Li, M., & Guo, B. (2019). Disrupters as well as monitors: Roles of others during and after collaborative remembering in DRM procedure. *Advances in Cognitive Psychology*, 15, 276–289. doi: 10.5709/acp-0275-1
- Osth, A. F., & Fox, J. (2019). Are associations formed across pairs? A test of learning by temporal contiguity in associative recognition. *Psychonomic Bulletin & Review*, 26, 1650–1656. doi: 10.3758/s13423-019-01616-7
- Peng, Y., Liu, Y., & Guo, C. (2019). Examining the neural mechanism behind testing effect with concrete and abstract words. *NeuroReport*, 30, 113–119. doi: 10.1097/WNR.0000000000001169
- Peterson, D. J., & Wissman, K. T. (2018). The testing effect and analogical problem-solving. *Memory*, 26, 1460–1466. doi: 10.1080/09658211.2018.1491603
- Rickard, T. C., & Pan, S. C. (2018). A dual memory theory of the testing effect. *Psychonomic Bulletin & Review*, 25, 847–869. doi: 10.3758/s13423-017-1298-4
- Schmitz, T. W., De Rosa, E., & Anderson, A. K. (2009). Opposing influences of affective state valence on visual cortical encoding. *Journal of Neuroscience*, 29, 7199–7207. doi: 10.1523/JNEUROSCI.5387-08.2009
- Siegel, A. L., Graup, R. S., & Castel, A. D. (2019). Emotion-enhanced binding of numerical information in younger and older adults. *Quarterly Journal of Experimental Psychology*, 73, 134–145. doi: 10.1177/1747021819867292
- Szöllösi, Á., Keresztes, A., Novák, B., Szászi, B., Kéri, S., & Racsmany, M. (2017). The testing effect is preserved in stressful final testing environment. *Applied Cognitive Psychology*, 31, 615–622. doi: 10.1002/acp.3363
- Verkoeijen, P. P. J. L., Bouwmeester, S., & Camp, G. (2012). A short-term testing effect in cross-language recognition. *Psychological Science*, 23, 567–571. doi: 10.1177/0956797611435132
- Ye, J., Nie, A., & Liu, S. (2019). How do word frequency and memory task influence directed forgetting: An ERP study. *International Journal of Psychophysiology*, 146, 157–172. doi: 10.1016/j.ijpsycho.2019.10.005
- Zimmerman, C. A., & Kelley, C. M. (2010). “I’ll remember this!” Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language*, 62, 240–253. doi: 10.1016/j.jml.2009.11.004
- Zhou, W., Nie, A., Xiao, Y., Liu, S., & Deng, C. (2020). Is color source retrieval sensitive to emotion? Electrophysiological evidence from old/new effects. *Acta Psychologica*, 210, 103156. doi: 10.1016/j.actpsy.2020.103156

RECEIVED 02.08.2020 | ACCEPTED 04.12.2020

## SUPPLEMENTARY MATERIAL

TABLE A1.

Data for Figures 5 and 6

Subject	Old-Practice			Old-restudy			Rearranged-Practice			Rearranged-restudy			Oldnew-Practice			Oldnew-restudy			New		
	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive
1	0.9	0.8	0.8	0.55	0.55	0.7	0.3	0.15	0.2	0	0.05	0.15	0.05	0	0.05	0	0.1	0	0.05	0	0
2	0.95	0.9	0.9	0.8	0.85	0.7	0.05	0.05	0	0.05	0.1	0	0	0	0	0	0	0	0	0	0
3	0.8	0.85	0.7	0.75	0.85	0.9	0.25	0.1	0.15	0.2	0.15	0.25	0.1	0	0.05	0.05	0	0.1	0	0	0
4	0.75	0.9	0.95	0.6	0.8	0.7	0	0.1	0	0	0.05	0.05	0.05	0	0	0	0	0	0	0	0
5	0.85	0.95	0.9	0.9	1	0.85	0.05	0	0.1	0.1	0	0.1	0.05	0	0	0	0	0	0	0	0
6	0.9	0.9	0.95	1	0.95	1	0.15	0.25	0.05	0	0	0	0.05	0	0	0	0	0	0	0	0
7	1	0.95	1	1	1	0.9	0.05	0	0.05	0	0	0.1	0	0	0	0	0	0	0	0	0
8	0.75	0.9	0.9	0.85	0.95	0.85	0.05	0	0	0	0	0.05	0	0	0	0	0	0	0	0	0
9	0.9	0.9	0.85	0.8	0.8	0.95	0.1	0.1	0.1	0.05	0.1	0.1	0	0.05	0	0.05	0	0	0	0	0
10	0.85	0.9	0.7	0.9	0.8	0.7	0.1	0.1	0.05	0.1	0.05	0.05	0	0	0.05	0	0	0	0	0	0
11	0.95	0.95	0.95	0.8	0.85	0.9	0.1	0.05	0.05	0.05	0	0.05	0	0	0	0	0	0.05	0	0.05	0
12	0.95	0.95	1	0.95	1	1	0.05	0.1	0.05	0.05	0.05	0.15	0.05	0	0	0	0	0	0	0	0
13	0.75	0.85	0.85	0.7	0.45	0.45	0.1	0.05	0.15	0.25	0.05	0.15	0	0	0	0	0	0	0	0	0
14	0.95	1	1	0.95	1	1	0	0	0	0.05	0	0.05	0.05	0	0	0	0	0	0	0	0
16	0.7	0.7	0.5	0.85	0.9	0.7	0.15	0.1	0.1	0.25	0.05	0.15	0.05	0	0	0.1	0.05	0	0	0	0
17	0.7	0.8	0.75	0.8	0.85	0.55	0.2	0	0.1	0.4	0	0.15	0.1	0	0	0	0	0	0.1	0	0
18	1	1	0.95	0.9	0.95	1	0.05	0.05	0	0.05	0.05	0	0	0	0	0	0	0	0.05	0	0
19	0.95	1	0.95	0.9	1	1	0	0	0.05	0	0	0	0	0	0	0	0	0	0	0	0
20	0.85	1	0.95	1	1	0.75	0	0.2	0.15	0.15	0	0.25	0	0.1	0.05	0.05	0	0	0	0	0
21	0.95	0.9	0.7	0.85	0.95	0.95	0.2	0.05	0.05	0.1	0.05	0.05	0	0	0.05	0	0	0	0	0	0
22	0.6	0.65	0.65	0.4	0.4	0.45	0	0.1	0	0	0	0.2	0	0	0	0	0	0	0	0	0
24	0.95	0.9	0.95	0.9	0.8	0.95	0.25	0.2	0.2	0.3	0.15	0.2	0	0	0	0.05	0	0	0	0	0
25	1	1	1	1	1	0.95	0.15	0.05	0.05	0.05	0.05	0.15	0	0.05	0	0	0	0	0	0	0
26	0.9	0.95	0.95	0.9	0.9	0.85	0.25	0.25	0.25	0.05	0	0.15	0	0.1	0	0	0	0	0	0	0
27	0.9	0.95	0.95	0.95	0.95	0.95	0	0.05	0.05	0	0	0	0	0	0	0	0	0	0	0	0
28	1	0.95	1	0.95	1	1	0	0	0.05	0	0	0.05	0	0	0	0	0	0.05	0	0	0
29	1	1	0.75	0.9	0.8	0.75	0	0	0.15	0	0	0.1	0.05	0	0	0	0	0	0	0	0
30	1	1	1	0.9	0.95	0.9	0	0	0.05	0.1	0	0.1	0	0	0	0.05	0	0	0.05	0	0
31	0.9	1	0.9	0.65	0.9	0.9	0.1	0	0.05	0	0	0	0	0	0	0	0.05	0	0	0	0
32	0.95	1	0.8	0.9	0.95	0.85	0.2	0.1	0.1	0.1	0	0.1	0.1	0	0	0	0	0	0	0	0