Age-related Changes in Working Memory: Age Affects Relational but not Conjunctive Feature Binding

Alexander Kirmsse, Saarland University

Hubert D. Zimmer, Saarland University & University of Western Australia

Ullrich K. H. Ecker, University of Western Australia

Department of Psychology, Brain and Cognition Unit, Saarland University; Hubert D. Zimmer, Department of Psychology, Brain and Cognition Unit, Saarland University & School of Psychological Science, University of Western Australia; Ullrich K. H. Ecker, School of Psychological Science, University of Western Australia.

This research was conducted within the International Research Training Group Adaptive Minds supported by German Research Foundation (DFG) under Grant IRTG 1457.

This research formed part of the first author’s PhD thesis at Saarland University. Parts of the results were presented at the Sixth International Conference on Memory (ICOM), Budapest, Hungary, 2016.

Correspondence concerning this article should be addressed to Alexander Kirmsse,
Department of Psychology, Saarland University, Campus A 2.4, D-66123 Saarbruecken,
Germany. E-Mail: alexander.kirmsse@uni-saarland.de
Abstract

In contrast to long-term memory, age-related association deficits in working memory are found only inconsistently. We hypothesized that type of binding is critical for the occurrence of such deficits. Relational binding abilities (associating separate visual units) should degrade with age, whereas more automatic conjunctive binding abilities (associating features within an object) should not. We contrasted associative memory and item memory using a change-detection task with colors and shapes in younger (18-33 years) and older (64-82 years) healthy adults. Color was either a surface feature of the shape (conjunctive binding) or a feature of a shape-external frame (relational binding). In a direct test of associative memory, participants memorized color-shape associations; in an indirect item memory test, participants were required to memorize only the shapes, and we measured the costs of ignoring task-irrelevant color changes from study to test. In the direct test, associative memory was poorer when relational binding was required rather than conjunctive binding, and associative memory was poorer in the older group, but no age-related association deficit was apparent. In the indirect test, by contrast, type of binding interacted with age: younger participants showed study-test congruence effects independent of the type of binding, but older adults showed enhanced congruence effects for conjunctive stimuli, indicating intact or even enhanced conjunctive binding, and practically no costs for relational stimuli, indicating poor relational binding. This stimulus-specific effect of a task-irrelevant feature change indicates that relational and conjunctive binding in working memory are differently affected by healthy aging.

**Keywords:** aging, working memory, binding, relational memory

**Word count (main text):** 7703 words
AGE-RELATED CHANGES TO RELATIONAL BINDING

Age-related Changes in Working Memory: Age Affects Relational but not Conjunctive Feature Binding

It is a well-established finding that memory decreases with advancing age, both with long (e.g., Craik & Rose, 2012) as well as short retention intervals (Park & Payer, 2006; Sander, Lindenberger, & Werkle-Bergner, 2012). Studies of episodic long-term memory have shown that elderly participants show a characteristic deficit in storing associative information, in particular when it comes to specific combinations of items or combinations of items and their contexts. By contrast, memory for individual items is relatively preserved (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008, for a meta-analysis). Interestingly, while such an age-related association deficit is consistently reported for long-term memory, a corresponding deficit in short-term or working memory (WM) has been reported only in some studies (Borg, Leroy, Favre, Laurent, & Thomas-Anterion, 2011; Brown & Brockmole, 2010; Chen & Naveh-Benjamin, 2012; Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Fandakova, Sander, Werkle-Bergner, & Shing, 2014; Mitchell, Johnson, Raye, Mather, & D'Esposito, 2000) but not others (Brockmole & Logie, 2013; Brockmole, Parra, Della Sala, & Logie, 2008; Brown & Brockmole, 2010; Brown, Niven, Logie, Rhodes, & Allen, 2017; Parra, Abrahams, Logie, & Della Sala, 2009; Rhodes, Parra, & Logie, 2016). The question thus arises: what is the factor that differentiates studies that have demonstrated an age-related association deficit in WM from studies that have not?

In long-term memory, the association deficit is typically assessed by pair recognition tasks. Participants study lists of items consisting of, for example, pairs of objects or objects in item-unique spatiotemporal contexts. Subsequently, they are required to recognize old pairs among recombined and/or new pairs (e.g., Naveh-Benjamin, 2000). However, to assess associative memory performance in WM, variants of change-detection paradigms are often used. In this paradigm, a study display is presented for a short amount of time, followed by a
retention interval of approximately one second. Then, participants decide whether a test probe (a single item or the whole display) is identical to or different from the study display. In this paradigm, the association usually refers to combinations of item features, for example the bindings of item shapes and their colors (e.g., when deciding whether the color of one of the items in the display has changed from study to test). In some tasks, however, the association refers to the bindings between items and their locations or contexts (e.g., when deciding whether the spatial item configuration has changed). Performance in these conditions is then compared to memory for the items themselves (for an overview, see Allen, 2015).

These procedural task differences—specifically, which type of binding is assessed—may be critical for the presence or absence of an age-related binding deficit in WM. In fact, studies that found no evidence for age-related association deficits have tended to focus on bindings between items and their surface features, whereas studies suggesting age-related association deficits have tended to focus on bindings between items and their locations or spatiotemporal contexts.

In long-term memory, these two different types of binding have been differentiated as mechanisms of conjunctive versus relational memory (e.g., Mayes, Montaldi, & Migo, 2007). Zimmer and colleagues have discussed them as intra-item versus relational binding (Zimmer, Mecklinger, & Lindenberger, 2006), or as intrinsic versus extrinsic binding (Ecker, Zimmer, & Groh-Bordin, 2007a, 2007b; Ecker, Zimmer, Groh-Bordin, & Mecklinger, 2007). Intra-item binding is considered to arise from a rather automatic perceptual process, which integrates the perceived features of an object. When this percept is consolidated into a memory representation, this integrated memory token can then be matched against a new perceptual input, which again is thought to be a rather automatic process (Treisman, 2006; Zimmer & Ecker, 2010). In contrast, relational binding refers to the ability to associate two independent tokens in perception and memory (see also Cabeza, 2006) and is assumed to be a
more deliberate process. Thus, conjunctive feature binding could be a byproduct of the encoding process provided by the perceptual stream, whereas relational binding could be a supplementary process that is only executed if the task context demands it. These two binding mechanisms may rely on different neural structures (see below) and, therefore, might undergo different changes with advancing age.

The theoretical distinction between relational and conjunctive binding in long-term memory is supported by electrophysiological findings (Ecker et al., 2007a, 2007b; Ecker, Groh-Bordin, & Zimmer, 2004; Yonelinas, 2002; Zimmer & Ecker, 2010) as well as brain imaging data (Cansino, Maquet, Dolan, & Rugg, 2002; Davachi, Mitchell, & Wagner, 2003). For example, in an event-related potentials study, Ecker et al. (2007b) presented participants with colored line drawings of everyday objects. Color was either an object feature (conjunction) or presented as a frame surrounding the object (relation). Task-irrelevant color changes influenced the frontal old-new effect associated with item familiarity only when color was an object feature but not if it was a feature of the relational frame. A comparable manipulation was used in an fMRI study by Staresina and Davachi (2009): relational but not conjunctive stimuli elicited greater hippocampal activity for later-remembered compared to later-forgotten shape-color associations.

The distinction between relational and conjunctive binding has more seldom been discussed in the context of WM (e.g., Allen, Brown, & Niven, 2013; Zimmer, 2008) but those WM studies that manipulated the type of binding required have supported the differentiation: Compared to relational binding, conjunctive binding has been found (1) to arise more involuntarily (Ecker, Maybery, & Zimmer, 2013), (2) to be less susceptible to interference (Allen, Hitch, Mate, & Baddeley, 2012; Delvenne & Bruyer, 2004; van Geldorp, Parra, & Kessels, 2015), (3) to be less affected by articulatory suppression (Walker & Cuthbert, 1998), and (4) to be associated with better associative memory performance (Delvenne & Bruyer,
Additionally, while relational binding has been associated with activity in the medial temporal lobes and the hippocampus (Hannula & Ranganath, 2008; Parra et al., 2015; Piekema, Kessels, Mars, Petersson, & Fernández, 2006; Piekema, Kessels, Rijpkema, & Fernández, 2009; Piekema, Rijpkema, Fernández, & Kessels, 2010), conjunctive binding has been linked to activity in parietal, occipital, and temporal brain regions (Parra, Della Sala, Logie, & Morcom, 2014; Piekema et al., 2006; Piekema et al., 2010; Shafritz, Gore, & Marois, 2002).

These differences suggest that the type of binding may also be relevant for understanding the influence of aging on WM performance. In particular, if relational but not conjunctive binding relies on the integrity of hippocampal structures (Staresina & Davachi, 2009) and these areas are found to be affected by age-related changes (Fjell & Walhovd, 2010; Raz et al., 2005; Raz, Ghisletta, Rodrigue, Kennedy, & Lindenberger, 2010; Shing et al., 2010), it follows that relational binding abilities should decline with age more strongly than conjunctive binding abilities.

As alluded to earlier, indirect evidence for differential effects of aging on relational versus conjunctive binding in WM comes from the comparison of studies utilizing conjunctive conditions with those utilizing relational conditions. Studies focusing on conjunctive binding have by-and-large failed to find age-related binding deficits in WM. For example, Brockmole and Logie (2013) investigated WM performance of more than 55,000 online participants from different age groups in a change-detection paradigm probing conjunctive shape-color binding. While older participants showed poorer item memory performance than younger participants, they demonstrated only minimal additional binding impairments. Similar results have been found in lab studies (Brockmole et al., 2008; Parra et al., 2009), even with extended encoding times (Brown et al., 2017; Rhodes et al., 2016) or addition of a secondary task (Brown & Brockmole, 2010; but see their Experiment 2).
Likewise Isella, Molteni, Mapelli, and Ferrarese (2015) reported only a small and non-significant age effect on shape-color binding. Overall, these findings imply that conjunctive binding is not substantially more affected by age than WM for items.

By contrast, age-related binding deficits have been demonstrated with WM tasks that arguably require relational integration, such as object-location binding with figural objects (Cowan et al., 2006; Mitchell, Johnson, Raye, & D’Esposito, 2000; Mitchell, Johnson, Raye, Mather et al., 2000), letters (Fandakova et al., 2014), emotional pictures (Borg et al., 2011; although binding was not contrasted with item memory in this study), and abstract shapes (Peich, Husain, & Bays, 2013). An age-related deficit in object-location binding seems to exist in WM—however, it might be present only under demanding procedural conditions. For example, when Rhodes, Parra, Cowan, and Logie (2017) excluded duplicate stimuli from the paradigm used by Cowan et al. (2006)—thus reducing the executive control necessary to prevent interference—they did not find an age-related binding deficit. Similarly, Read, Rogers, and Wilson (2016) did not find an age-related binding deficit when they controlled for interference from multi-stimulus probes. Moreover, as pointed out by Rhodes et al. (2017), the reported evidence may be weaker than it first appears because in some studies (e.g., Borg et al., 2011; Fandakova et al., 2014; Mitchell, Johnson, Raye, Mather et al., 2000), only significant vs. non-significant simple main effects were reported but not a significant interaction effect (cf. Nieuwenhuis, Forstmann, & Wagenmakers, 2011). With regard to object-context binding, Chen and Naveh-Benjamin (2012) reported a binding deficit in older adults if participants were required to memorize the associations between faces and scenes presented as backgrounds.

To the best of our knowledge, only three studies (Bastin, 2017; Peterson & Naveh-Benjamin, 2016; van Geldorp et al., 2015) have directly compared the age-sensitivity of conjunctive and relational binding in WM within the same experiment. Bastin (2017) created
a WM task involving real-world object line drawings with color being either a surface feature (e.g., a yellow tent) or a relational context feature (i.e., the tent was presented as a black-and-white line drawing and color as a patch next to the tent). At test, participants were required to decide whether a presented object-color association was part of the study list. Compared to WM performance for single features, older but not younger adults had a disproportionately lower WM performance for relational compared to conjunctive object-color associations.

Van Geldorp and colleagues (2015) tested younger and older participants in a WM task designed to investigate relational and conjunctive shape-color binding. The task involved reconstructing study items by assigning shapes to their study color, and was performed with or without a concurrent interfering task (counting backwards). Supporting the distinction between relational and conjunctive binding, overall memory performance was lower for relational than for conjunctive stimuli, and the interfering task impaired memory for relational items more than memory for conjunctive items. Notably, however, while older participants showed poorer memory performance than younger participants, there were no interactions involving age and type of binding even though relational binding was more demanding than conjunctive binding.

Peterson and Naveh-Benjamin (2016, Experiment 1) used a change-detection task requiring memory for items—colors or shapes—and contrasted this with a change-detection task requiring memory for shape-color bindings. Color was either a surface feature of the shape (which was presented on a gray background square) or it was a background feature (i.e., the shape was gray and presented on a colored square). The type of binding had no effect on performance. However, Peterson and Naveh-Benjamin observed an age-related association deficit: Older people showed a larger binding deficit than younger participants. This age-related binding deficit disappeared when participants engaged in articulatory suppression during study, leaving only a main effect of age.
The hitherto available data allow three conclusions regarding age-related effects in WM: (1) conjunctive information is processed rather automatically across age groups, without any age-related impairments; (2) in some studies, relational object-location binding was negatively affected by aging if distractor competition was high; and (3) evidence for age-related impairments in object-context binding is sparse and inconsistent. One reason for the diverging results regarding relational binding might lie in the explicit binding requirements of the change-detection paradigm used in the majority of studies. That is, participants are typically instructed to attend to and encode feature combinations for a test that directly targets the bindings. With such a test procedure, intentional use of encoding strategies is likely (see Peterson & Naveh-Benjamin, 2016). Hence, it is possible that older people have an encoding deficit for relational bindings, which, however, they may be able to compensate for by enhanced intentional effort.

One way to circumvent speculations regarding the impact of potentially obscuring variables such as strategy use and effort is to utilize indirect binding tests. Such indirect tests purport to be item memory tests; participants are instructed to focus on and memorize items (e.g., object shapes) while ignoring other features of the presented study objects (e.g., object color or background color). To the extent that feature information is automatically integrated in the WM representation, study-test changes of the task-irrelevant feature should impair item memory (e.g., slow down recognition; Gao, Gao, Li, Sun, & Shen, 2011). This paradigm has been used to demonstrate the obligatory nature of conjunctive but not relational binding both in long-term memory (Ecker et al., 2007a, 2007b; Nicholson & Humphrey, 2004; Zimmer & Ecker, 2010) and WM (Ecker et al., 2013; Maybery et al., 2009) in young adults. However, no studies to date have used indirect tests to investigate WM binding in older participants.

Therefore, the goal of the present study was to contrast short-term memory for relational and conjunctive stimuli in young and older adults using both direct and indirect
tests. Based on the existing literature outlined above, we hypothesized that memory for associations of relational features should generally be poorer than memory for associations of conjunctive features. Also, no age effects were expected for conjunctive stimuli. By contrast, to the extent that relational binding is affected by age, older adults should show a particularly strong impairment of memory for relational associations, resulting in an interaction between type of binding and age.

The exact difference between item and binding conditions depends on the degree of independence between memory for individual features and memory for bindings. In the literature, for conjunctive items, three relationships between feature memory and memory for bindings (i.e., in the conjunctive case: memory for objects) have been reported. First, some have argued for an all or none relationship. For example, in the study from Luck and Vogel (1997), participants remembered all features of an object, or none. Thus, feature memory was completely dependent on object memory, and performance in the binding condition was at the level of feature memory performance. Second, others have hypothesized that feature memory is object-based but independent, progressing from coarse to fine (e.g., Gao, Ding, Yang, Liang, & Shui, 2013). Based on this, one would expect that performance in the binding condition is at a similar level as performance for the most difficult-to-remember individual feature. Indeed, this result has been obtained in a number of studies (e.g., Kursawe & Zimmer, 2015; Song & Jiang, 2006; Woodman & Vogel, 2008). Finally, sometimes memory for different features has been found to be independent (e.g., Bays, Wu, & Husain, 2011). In this case, memory for bindings is expected to be equal to the product of the recognition probabilities for the individual features, as long as memory is not positively influenced by additional binding processes or negatively affected by binding deficits.

Irrespective of the exact nature of the relationship between feature memory and binding memory, an interaction effect was expected in the present study: If older adults show
a specific binding deficit, any decrement in binding memory compared to feature memory should be magnified relative to younger participants, causing an interaction between the change condition (feature vs. binding change) and age. However, if older people can invest additional effort to compensate for their relational-association deficit when they are explicitly instructed to memorize the associations, then the interaction should be moderated by the type of test. That is, the proposed age-related relational binding deficit should be evident in the indirect test but may not occur with a direct test (Note that binding deficits will present as reduced memory performance for item-feature associations in the direct test, and reduced costs associated with task-irrelevant study-test feature changes in the indirect item recognition test).

In order to test these hypotheses, younger and older participants completed a change-detection task. Participants studied sets of colored shapes (conjunctive stimuli) or white shapes encased in colored frames (relational stimuli). Participants completed the task under two conditions: in the direct association-test condition, they were instructed to memorize the color-shape bindings; in the indirect test condition, they memorized only the item shapes. The single test probe was either identical to a study item (no change), it contained an unstudied shape or color (new shape, new color), or it was a new arrangement of a studied color and a studied shape (recombination). Test probes with a new color or a new shape allowed an assessment of item memory, as responses in these conditions could be made based on the evaluation of individual features rather than the specific shape-color association. By contrast, test probes recombining old features allowed an assessment of binding performance, as knowledge for the specific study associations was required for accurate performance (see Olson & Jiang, 2002; Parra et al., 2009; Wheeler & Treisman, 2002).

Using a single test probe is recommended if the test involves recombinations (Cowan, Blume, & Saults, 2013), and a single test probe is also standard for indirect tests. To keep
conditions comparable, a single test probe was thus used in both the direct and indirect tests. The probe was presented in a central position to minimize influences of object-location binding on memory (see van Geldorp et al., 2015).

**Method**

The experiment used a fully-crossed 2 (test type: direct, indirect) × 2 (stimulus type: relation, conjunction) × 4 (change type: same, new shape, new color, recombined) × 2 (age: young, old) mixed design, with age as the only between-subjects factor.

**Participants**

Sample size was determined by an a-priori power analysis using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). To this end, we initially obtained effect-size estimates for the critical interaction between age and type of binding from previous studies reporting an association deficit, using the error terms of the analyses of variances (ANOVA). The obtained effect sizes were mostly $\eta_p^2 > .18$. Based on a conservative estimate of $\eta_p^2 = .15$, the suggested minimum sample size to detect an interaction effect with power $1 - \beta = .95$ was 38 participants per group.

A total of 93 participants took part in the study; 48 were young adults—non-psychology undergraduate students from Saarland University—and 45 were older adults. Two young and two older participants reported problems understanding the instructions, and one young participant confused the response buttons. These participants were excluded, so the final sample size was $N = 88$, comprising $n = 45$ young adults (age range 18-33 years; mean age $M = 21.93$, $SD = 3.28$; 23 females, 22 males) and $n = 43$ older adults (age range 64-82 years; mean age $M = 71.33$, $SD = 4.61$; 30 females, 13 males).

Participants’ processing speed was assessed with a digit symbol substitution test adapted from Wechsler (2008). Young participants’ mean score was $M = 64.67$ ($SD = 11.65$), whereas older participants’ mean was $M = 47.65$ ($SD = 9.84$). Both values are in a plausible
range (see Ferdinand & Kray, 2013; Kray, Eber, & Karbach, 2008). The younger sample had more years of formal education ($Md = 13, M = 12.46, SD = 0.77$) than the older sample ($Md = 10, M = 10.72, SD = 1.65$), $U = 352.50, z = -4.47, p < .001$.

All participants reported good health, normal or corrected-to-normal vision, no history of neurological disorders (in the last 5 years), no deficits of hand mobility, and no use of medication that might influence response times. All participants provided written informed consent after reading an ethically approved information sheet; they received a monetary compensation of €8 per hour. The study was performed under ethical approval of the Ethics Committee of the Philosophical Faculty III Empirical Human Sciences at Saarland University.

Materials and Apparatus

Stimuli consisted of 24 novel complex shapes for which no common names exist. For this purpose, different simple geometric shapes were overlaid and their outline defined the border of the complex figure (see Figure 1 for examples). Two versions of the stimuli were generated: conjunction stimuli were fully colored shapes; relation stimuli were white shapes (RGB values: 255, 255, 255) encased with a colored rectangular frame. The number of colored pixels in the frame approximated the number of pixels contained in the fully colored version of the shape, in order to achieve physically similar amounts of color information. Ten different colors were used: red (255, 0, 0); green (0, 255, 0); blue (0, 0, 255); magenta (255, 0, 255); cyan (0, 255, 255); yellow (255, 255, 0); orange (255, 102, 0); brown (153, 102, 51); pink (255, 153, 204); and black (0, 0, 0). The shape-stimuli subtended $1.98^\circ \times 1.98^\circ$ of visual angle, the surrounding frame $3.80^\circ \times 3.80^\circ$. Stimuli were presented on a gray (153, 153, 153) background. Stimuli were presented on a 23-inch flat-screen monitor with a resolution of $1,920 \times 1,080$ pixels and a refresh rate of 60 Hz. Participants were seated in front of the screen at an approximate viewing distance of 60 cm.
To ensure comparisons truly reflect differences in memory performance and not potential differences in sensory-perceptual abilities between older and younger participants, presentation time of study items was adjusted for each participant in order to compensate for age-related slowing of encoding. It has been shown that sensory deficits can partially account for observed differences in the associative memory measures and other cognitive deficits (Baltes & Lindenberger, 1997; Lindenberger & Baltes, 1994; Naveh-Benjamin & Kilb, 2014). Even for younger adults, insufficient encoding time can reduce WM performance (Eng, Chen, & Jiang, 2005). To calculate the necessary adjustment for each participant, a visual search task (see Alvarez & Cavanagh, 2004; Vaughan & Hartman, 2010), adapted to estimate the individual encoding times for the main experiment, was administered prior to the actual WM task (see Appendix for details). It should be noted that this adjustment only served to equate perceptual encoding across groups, without influencing memory per se; this is because longer presentation times do not enhance memory if the modulation only serves to approximate the time necessary for proper encoding (e.g., Luck & Vogel, 1997; Rhodes et al., 2016; Sander, Werkle-Bergner, & Lindenberger, 2011; Vaughan & Hartman, 2010).

Procedure

The trial structure is illustrated in Figure 1. Each trial began with the presentation of a central fixation cross. After 600-800 ms (the exact interval was randomly determined), a cue word was presented centrally, just above the fixation cross, for 800 ms. The cue word indicated the task-relevant item information: the cue “BEIDES” (both) was used in trials of the direct test; the cue “FORM” (shape) was used in trials of the indirect test. The fixation cross remained on the screen for another randomly chosen interval between 750-1,050 ms after the cue word disappeared. Next, an array of three study stimuli was presented; stimuli were arranged on an invisible circle (6° visual angle in diameter) centered on the midpoint of the screen, with the three item locations forming a virtual isosceles triangle. The presentation
time of the study display was on average 1,493 ms ($SD = 293$ ms) for younger, and 1,846 ms ($SD = 291$ ms) for older adults, individually adjusting for each participant’s encoding speed as estimated from the visual search task (see Appendix for details). These presentation times are in the range of other WM studies with older samples (e.g., Vaughan & Hartman, 2010).

After a retention interval of 1,000 ms, a single probe was presented in the center of the screen. On direct test trials (shape-color associations are task-relevant), participants indicated whether the presented shape-color combination had featured in the study display or not. On indirect test trials (new color and color recombinations are task-irrelevant), they indicated whether or not an item of the same shape had been presented during study. Participants responded by pressing a key on a Cedrus response pad (RB-834, Cedrus Corporation). Response categories (match and non-match) were assigned according to participants’ handedness, with match trials mapped onto the dominant hand (in case of uncertainty, handedness was assessed according to the Edinburgh inventory; Oldfield, 1971). The test display remained on screen until a response was given or for a maximum of 5,000 ms; it was followed by an inter-trial interval of 1,200 ms. Both accuracy and response speed were emphasized.

There were a total of 360 trials; half of these (180) used conjunctive stimuli and half relational stimuli. For each stimulus type, half the trials (90) were direct association test trials, and half were indirect test trials. In each test condition, half the trials (45) were change trials and half required a no-change response. Thus, in the direct test, 15 trials occurred in each change condition: new color, new shape, and recombination. In the indirect test, no-change stimuli, stimuli with a new color, and stimuli with a recombinated color (15 trials each) required a no-change response and 45 trials fell in the new shape category. Trials were in random order. To ensure adequate understanding of the paradigm, there were 24 practice trials that were not included in analyses.
Results

Trials with response times falling 2.2 times the interquartile range above the third quartile or below the first quartile of a participant’s individual response time distribution (Hoaglin & Iglewicz, 1987; Hoaglin, Iglewicz, & Tukey, 1986), as well as trials with response times below 200 ms were excluded prior to all analyses. This resulted in the exclusion of 2.55% of data and reduced the positive skew of the data. Nevertheless, response times were still non-normally distributed; we therefore based analysis on log10-transformed RT data. However, as results did not differ qualitatively and for ease of comprehension, we report descriptives based on the untransformed RT data.

Dependent variables were recognition accuracy and response times. Corrected recognition scores were calculated as $Pr = \text{hit rate} - \text{false alarm rate}$ (i.e., proportion of hits + proportion of correct rejections − 1; Snodgrass & Corwin, 1988). The $Pr$ measure is associated with a two high-threshold model, which we considered plausible for the type of task administered: Participants can explicitly judge the match (or mismatch) of the test probe and their memory representation and guess if there is insufficient evidence for either a “different” or “same” response. As direct and indirect tests required different responses, scoring and analyses were conducted separately (see details below). Follow-up analyses were obtained by investigating contrasts or (Bonferroni-corrected) pairwise comparisons. In all figures, we present 95% confidence intervals based on within-subjects variance to make a direct comparison of conditions possible (following Cousineau, 2005). In the text, we present means together with their 95% confidence intervals (in squared brackets) based on between participants variance in order to give an impression of the variability of data.
Direct Association Test

In the direct association test, a “hit” was defined as the correct detection of a change; a “false alarm” was defined as incorrect change detection in no-change trials (i.e., same condition). Table 1 shows mean $Pr$ scores across conditions dependent on age.

Performance in the new shape and new color conditions was averaged to obtain a composite index of item memory. Item memory was contrasted with performance in the recombination condition—which provided an index of binding memory—in a $2 \times 2 \times 2$ mixed-measures analysis of variance (ANOVA) with within-subjects factors stimulus type (relation, conjunction) and change type (item [i.e., new shape/new color], binding [i.e., recombination]), and the between subjects factor age (young, old). Results are displayed in Figure 2.

The analysis yielded a main effect of stimulus type: memory performance was greater for conjunctive stimuli, $M = .57$, 95% CI [.54, .59], $SE = .01$, compared to relational stimuli, $M = .44$, 95% CI [.41, .47], $SE = .015$, $F(1, 86) = 71.57, p < .001, \eta_p^2 = .45$. There was also a main effect of change type: detection of shape-color recombinations, $M = .41$, 95% CI [.38, .44], $SE = .016$, was lower than detection of item changes (i.e., new shape or new color), $M = .59$, 95% CI [.57, .61], $SE = .011$, $F(1, 86) = 215.96, p < .001, \eta_p^2 = .72$. These main effects were qualified by a marginal interaction between stimulus type and change type, $F(1, 86) = 3.28, p = .074, \eta_p^2 = .04$. The main effect of age was also significant: older participants, $M = .44$, 95% CI [.41, .48], $SE = .017$, performed worse than younger participants, $M = .56$, 95% CI [.53, .59], $SE = .017$, $F(1, 86) = 23.17, p < .001, \eta_p^2 = .21$. There were no two-way interactions involving age (all $F$s < 1), and the three-way interaction was also non-significant, $F(1, 86) = 1.99, p = .16, \eta_p^2 = .02$. In sum, results demonstrated generally reduced recognition performance for relational (as compared to conjunctive) stimuli.
and a general impairment with age. There was, however, no indication of an age-related binding deficit in WM, neither for conjunctive nor for relational stimuli.\(^1\)

Additionally, we analyzed item memory separately for shape and color change conditions (see Figure 3). To this end, we ran a 2 × 2 × 2 mixed-measures ANOVA with within-subjects factors stimulus type (relation, conjunction) and change type (new shape, new color), and the between-subjects factor age (young, old). Apart from effects already demonstrated in the previous analysis, this analysis yielded a main effect of change type: new color changes, \(M = 0.63, 95\% \text{ CI} [.60, .66], SE = .015\), were detected more often than new shape changes, \(M = 0.55, 95\% \text{ CI} [.53, .57], SE = .012, F(1, 86) = 28.77, p < .001, \eta^2_p = .25\). However, this main effect was qualified by a three-way interaction with stimulus type and age, \(F(1, 86) = 4.53, p = .036, \eta^2_p = .05\). For conjunctive stimuli, the change-type effect was influenced by age, resulting in a significant interaction contrast of change type and age, \(F(1, 86) = 7.06, p = .009, \eta^2_p = .08\). No such effect was found for relational stimuli (\(F < 1\)). With conjunctive stimuli, relative to younger adults, older adults recognized color changes (i.e., an old shape in a new color) more often than expected and shape changes (i.e., a new shape in an old color) less often.

**Indirect Test**

In the indirect test, a “hit” was defined as the correct acceptance of a studied shape as old; a “false alarm” was defined as an “old” response to a new shape. Hit rates were calculated separately for the three different types of shape-matching trials (same, new color, recombined). Following ample precedent (e.g., Ecker et al., 2013; Gao et al., 2011), a secondary dependent variable in the analysis of the indirect test data was response time costs—the difference between hit response times to unchanged same items and old items with

---

\(^1\) To exclude the possibility that different trial numbers across feature and recombination conditions influenced the results, we repeated the analysis with change type as a three-level factor (new color, new shape, recombination); in this analysis, all conditions have the same number of trials. The analysis revealed the same two main effects as the combined analysis.
irrelevant study-test changes. Table 2 lists mean Pr scores across conditions dependent on age as well as the corresponding response time costs.

**Accuracy.** Pr scores were subjected to a 2 × 2 × 3 mixed-measures ANOVA with the within-subjects factors stimulus type (relation, conjunction) and change type (same, new color, recombin ed), and the between-subjects factor age (young, old) (see Figure 4). There were main effects of stimulus type, $F(1, 86) = 8.49, p = .005, \eta^2_p = .09$, change type, $F(2, 172) = 44.11, p < .001, \eta^2_p = .34$, and age, $F(1, 86) = 25.44, p < .001, \eta^2_p = .23$. These indicated that recognition performance was better for relational stimuli, $M = .63$, 95% CI [.60, .66], $SE = .014$, than conjunctive stimuli, $M = .59$, 95% CI [.56, .62], $SE = .016$, and better for same items, $M = .67$, 95% CI [.65, .70], $SE = .014$, compared to new color changes, $M = .57$, 95% CI [.53, .60], $SE = .018$, and recombined changes, $M = .59$, 95% CI [.56, .62], $SE = .014$; the latter two were not different from each other, $p = .15$. The size of this irrelevant-change effect varied with age and stimulus type, but the corresponding pattern was observed in all simple main effects. Again, young adults, $M = .68$, 95% CI [.64, .72], $SE = .019$, generally performed better than older adults, $M = .54$, 95% CI [.50, .58], $SE = .020$.

Notably, there was a significant interaction between stimulus type and change type, $F(2, 172) = 5.24, p = .006, \eta^2_p = .06$, which was further influenced by age, as demonstrated by a significant three-way interaction, $F(2, 172) = 3.44, p = .034, \eta^2_p = .04$. To better understand this interaction, we analyzed the data of the two age groups separately, in two 2 × 2 repeated measures ANOVAs with the factors stimulus type and change type. For young adults, there was only a main effect of change type, $F(2, 88) = 27.91, p < .001, \eta^2_p = .39$ (all other $Fs < 1$). However, the corresponding analysis in older adults yielded two significant main effects (stimulus type: $F(1, 42) = 10.43, p = .002, \eta^2_p = .20$; change type: $F(2.84) = 18.05, p < .001, \eta^2_p = .30$), as well as a significant interaction, $F(2, 84) = 7.16,$
AGE-RELATED CHANGES TO RELATIONAL BINDING

$p = .001, \eta_p^2 = .15$. Like young adults, older adults did not show a stimulus type effect for same relational, $M = .60$, 95% CI [.55, .65], $SE = .026$, and conjunctive items, $M = .60$, 95% CI [.55, .65], $SE = .023$. They did, however, show a stimulus type effect for items with task-irrelevant changes, $F(1, 42) = 19.77, p < .001, \eta_p^2 = .32$: the average recognition performance was lower for conjunctive stimuli with new color changes, $M = .44$, 95% CI [.38, .51], $SE = .033$, and recombination changes, $M = .49$, 95% CI [.44, .54], $SE = .026$, compared to relational stimuli with new color changes, $M = .55$, 95% CI [.50, .60], $SE = .026$, and recombination changes, $M = .57$, 95% CI [.53, .61], $SE = .021$.

In other words, the costs of task-irrelevant change differed across the two age groups. To illustrate, Figure 5 plots change costs, calculated as the difference between $Pr$ scores for unchanged same items minus $Pr$ scores for old shapes with an irrelevant color change (averaged across new-color and recombined conditions). Change costs were roughly equivalent for both stimulus types in young adults; yet, in older adults, change costs were sizeable for conjunctive stimuli but small and nearly insignificant for relational stimuli.

**Response time costs.** A pattern comparable to the results when analyzing $Pr$ scores was found for response times (RT). Longer RTs were expected in irrelevant-change trials due to the conflict that arises from the mismatch between test probe and memory representation (e.g., Ecker et al., 2013; Gao et al., 2011). In order to estimate RT change costs, RTs of the same condition were subtracted from RTs of the two irrelevant-change conditions (i.e., positive values represent costs). The log-transformed RT data were analyzed in a $2 \times 2 \times 2$ ANOVA with within-subjects factors stimulus type (relation, conjunction) and change type (new color, recombined), and the between-subjects factor age (young, old). A main effect of stimulus type was observed, $F(1, 86) = 5.20, p = .025, \eta_p^2 = .06$: RT costs were greater for conjunctive stimuli, $M = 107$ ms, 95% CI [66, 148], $SE = 21$, than for relational stimuli, $M = 47$ ms, 95% CI [14, 80], $SE = 17$. Although there was no significant main effect of age, $F < 1$,
there was a significant two-way interaction of stimulus type and age, \( F(1, 86) = 6.47, p = .013, \eta_p^2 = .07 \). Comparable to the accuracy analysis, pairwise comparisons indicated that RT costs did not differ between relational, \( M = 58 \text{ ms}, 95\% \text{ CI} [12, 105], SE = 23 \), and conjunctive stimuli, \( M = 64 \text{ ms}, 95\% \text{ CI} [7, 121], SE = 29 \), in younger adults, \( p = .85, d_{Cohen} = 0.03 \). In contrast, older adults showed significantly greater change costs with conjunctive stimuli, \( M = 150 \text{ ms}, 95\% \text{ CI} [91, 208], SE = 29 \), compared to relational stimuli, \( M = 36 \text{ ms}, 95\% \text{ CI} [-11, 84], SE = 24 \), \( p = .001, d_{Cohen} = 0.50 \).

**Discussion**

The goal of the present study was twofold. The first goal was to investigate whether age has a differential influence on the intentional storage and retrieval of relational and conjunctive feature bindings in WM. Binding performance was assessed in a direct test using shape-color associations in a change-detection paradigm. It was assumed that an age-related binding deficit would be present for relational feature associations. As expected, memory for bound relational information was generally poorer than memory for bound conjunctive information. Moreover, older participants showed decreased WM performance overall. However, contrary to expectations, there was no indication of any age-related binding deficit—neither in a general sense, nor specifically for relational information.

The second goal was to investigate whether age affects the involuntary binding of relational vs. conjunctive features. Age-related binding deficits were assessed in an indirect test where the association between features was not task relevant (i.e., an irrelevant-change paradigm). Compared to young adults, older adults were expected to show smaller costs associated with an irrelevant color change from study to test with relational items but not with conjunctive items. The results clearly supported this second hypothesis.
Processing of Relational and Conjunctive Information in WM

Both direct and indirect tests were influenced by the type of binding, demonstrating that the distinction between conjunctive and relational binding is critical not only in long-term memory but also in WM (see Ecker et al., 2013). Specifically, in the direct test, memory performance for conjunctive stimuli was higher than for relational stimuli, as has been observed in previous work (e.g., Delvenne & Bruyer, 2004; van Geldorp et al., 2015; Walker & Cuthbert, 1998; Xu, 2002b). Interestingly, the same effect was seen in our composite index of item memory, which suggests that the advantage is not solely due to a difference in the ease of binding—it seemed generally harder to memorize features of relational than conjunctive stimuli. This effect, however, was confined to the direct test: the processing advantage for conjunctive stimuli was not seen in the indirect test, where only shape information was relevant. Hence, conjunctive stimuli had an advantage over relational stimuli only when both parts of the relational stimuli (i.e., both the shape and the colored frame) were task relevant and had to be attended. The poorer memory performance associated with the relational items is probably not a consequence of a more difficult perceptual discrimination (i.e., crowding; e.g., Whitney & Levi, 2011). It is more likely the consequence of an “object benefit”: it is easier to attend to two features of the same object than attend to two features of disconnected objects (Fougnie, Cormiea, & Alvarez, 2013; Xu, 2002b, 2002a; Xu & Chun, 2007). This has consequences for the involuntary encoding of features, especially for elderly people, as will be discussed later.

Furthermore, in the direct test, performance in the binding condition (i.e., recombination) was generally poorer than in the item conditions (i.e., new shape/new color). Binding performance was also poorer than performance for the most difficult individual feature condition (i.e., new shape). This speaks against the notion that complete objects are represented in memory, and also against the notion that feature memory is object-based with
a coarse-to-fine gradient (in which case color should have always been co-represented with shape). By contrast, the results support the assumption that features were stored separately, and that items were incompletely represented in WM (Bays et al., 2011; Cowan et al., 2013; Fougnie et al., 2013; Kursawe & Zimmer, 2015). Accordingly, recognition performance in the recombined condition was only marginally (though statistically significantly\(^2\); average \(\Delta = .05\)) better than expected based on independent feature memory (i.e., the product of the recognition probabilities for the individual features). Thus, it is possible that for the majority of items, the probability of remembering both features equates to the joint probability of sampling color and shape of the same object (Vul & Rich, 2010). Considering this, even in the binding condition, WM performance may be mainly a consequence of remembering individual features, with only a small contribution of any supplementary binding mechanism. Surprisingly, this was the case in both age groups and with both types of items. We only got a main effect of the type of item, indicating that the proportion of items for which both features were remembered was clearly higher in the conjunction than the relation condition.

**Aging and WM Performance**

The present study replicates previous findings of a general decrease in WM performance with age. Such age effects are frequently reported across a wide range of different stimuli and tasks (for an overview, see Allen et al., 2013). Causal factors that have been proposed include less efficient executive control processes as well as less distinct

\(^2\) We calculated the multiplicative combination of the single-feature probabilities of new-color and new-shape changes for relational and conjunctive stimuli, and compared these values with the observed values in the binding conditions. A \(2 \times 2 \times 2\) ANOVA with the within-subjects factors stimulus type (relation, conjunction), and value (expected, observed) and the between-subjects factor age group (young, old) yielded the already known main effects of age and stimulus type; additionally, we found that observed values, \(M = .41, 95\%\) CI [.38, .44], \(SE = .015\), were on average higher than the expected values, \(M = .36, 95\%\) CI [.33, .39], \(SE = .013\), predicted from the independence assumption, \(F(1, 86) = 18.51, p < .001, \eta_p^2 = .18\). All interaction were non-significant (stimulus \(\times\) group, \(F(1, 86) = 1.57, p = .21, \eta_p^2 = .02\); stimulus type \(\times\) value, \(F(1, 86) = 1.37, p = .25, \eta_p^2 = .02\); value \(\times\) group, \(F < 1\); and stimulus type \(\times\) value \(\times\) group with \(F < 1\).
sensory representations in older adults (Gazzaley & Nobre, 2012). Interestingly, the results from the direct test in the present study suggest that the age-related memory impairment was present for item memory and binding conditions. As both conditions had different task demands, WM impairments in the older age group seemed independent of task demands. Speculatively, it could be argued that such deficits are most plausibly explained by fluctuations in attentional control (Adam, Mance, Fukuda, & Vogel, 2015), which can lead to items being “missed” independent of task difficulty.

**Aging and Processing of Conjunctive Information**

The absence of an age-related binding deficit for conjunction items was expected and replicates results from previous studies (Brockmole et al., 2008; Brockmole & Logie, 2013; Parra et al., 2009; Peterson & Naveh-Benjamin, 2016; van Geldorp et al., 2015). While the results from the direct test may not be considered strong evidence for the integrative processing of conjunctive information, more convincing evidence for integrative processing of conjunctive information is provided by the indirect test. These findings corroborate the notion that, independent of age, binding of a shape and its intrinsic surface color is a rather involuntary process (Ecker et al., 2013; Zimmer & Ecker, 2010).

In fact, the RT costs associated with the response conflict arising from irrelevant conjunctive feature changes were approximately twice as large in older compared to younger participants. Importantly, older participants showed comparable costs for recombined and new feature changes. This demonstrates that the slower responses were not an unspecific interference effect but a consequence of remembering the “correct” (i.e., studied) color. This suggests that in conjunctive stimuli, color was in fact bound to shape even though it was not task-relevant. In other words, a recombination of old features could only create conflict if the studied shape-color association is represented in memory. Larger RT costs for older participants are consistent with previous findings of greater age-related costs associated with
the filtering out of irrelevant information (Gazzaley, Cooney, Rissman, & D'Esposito, 2005; Hasher, Quig, & May, 1997; Jost, Bryck, Vogel, & Mayr, 2011; Oberauer, 2005). To conclude, the present study found no indication of an age-related binding deficit in WM for conjunctive stimuli; on the contrary, the data even suggest enhanced memory of irrelevant conjunctive information.

Aging and Processing of Relational Information

For relational stimuli, no age-related binding deficit was found in the direct test. This is an unexpected result that is inconsistent with some previous research (Bastin, 2017; Chen & Naveh-Benjamin, 2012; Walker & Cuthbert, 1998). The result is, however, in line with findings from recent studies that utilized a similar type-of-binding manipulation as the present study (Peterson & Naveh-Benjamin, 2016; van Geldorp et al., 2015). It seems that the spatially disjunct presentation of shape and color information in itself is not sufficient to establish an age-related association deficit in WM. Potentially, an age-related binding deficit may become apparent only when perceptual integration of the to-be-bound information is not possible even if participants know that memory of the combination is requested. For example, an age-related association deficit was found by Chen and Naveh-Benjamin (2012) with associations between faces and scenes. Faces and scenes belong to completely different domains (Mayes et al., 2007) and a face-scene binding task should therefore require additional processes for the association of these two perceptual units. This might have not been the case for the relational stimuli in the present study. We have already discussed that the independent encoding of color and shape from the same object can almost completely explain performance in the binding condition with both types of items. If this is correct, there is no supplementary binding process except for the allocation of attention to both features; this would make plausible the absence of an age-related binding deficit. Even older people should be able to perceptually group a shape and its surrounding frame if they intentionally
attend to both parts of the item (for a discussion, see for example Xu, 2002b), even though this will be harder than with conjunction items.

However, in contrast to the direct test, the indirect test results showed that the two types of stimuli were processed differently and that the two age groups differed in their propensity to bind relational information. In irrelevant-change conditions, a clear interaction between age and stimulus type emerged. In young participants, task-irrelevant feature changes produced equivalent performance costs with both relational and conjunctive items, suggesting that the irrelevant color was represented in WM independent of item type. By contrast, older participants showed significant change costs only for conjunctive stimuli, and costs were virtually absent for relational stimuli. Given the observation that older participants showed even larger change costs for conjunctive items, we interpret the absence of this effect for relational stimuli as a clear indication of an age-related change in relational binding.

Conclusions and Caveats

Our study provided several clear results: (1) The distinction between conjunctive and relational items is important for the understanding of binding. Relational items were remembered less well than conjunctive items, both in terms of feature memory and binding memory. (2) Relational items were differently processed by younger and older adults if the relational feature was task-irrelevant. Both groups integrated irrelevant conjunctive information, but only young participants integrated irrelevant relational information; the older participants did not do so. (3) An age-related binding deficit existed neither for spatially integrated (conjunctive) nor spatially separated (relational) shape and color information if the combination was intentionally encoded. (4) However, a general age-related memory deficit was observed despite our attempts to compensate for general age effects by adjusting encoding times.
The different integration of relational and conjunctive information could be the result of age-related changes to early selection mechanisms when visual information is encoded into WM, or the result of changes to later binding processes when information is consolidated in WM. The paradigm utilized in the present study was not designed to distinguish between these possibilities. For this purpose, other paradigms are necessary that are sensitive to online effects of processing. In total, however, the data suggest that the age-related differences in shape-color binding were due to selection effects. In the direct test, the item and age effects were visible in feature memory as well as in binding memory, and the effects were similar for both types of items. We have no evidence for a substantial binding process other than the joint encoding of color and shape features (see also Vul & Rich, 2010). The same mechanism might explain the difference in the indirect test: If younger but not older participants attended to the irrelevant relational color information, only the younger participants would represent the surrounding color in the object file representing an item. Cowan and colleagues have suggested that unitized objects and feature representations can coexist in WM and that the task demands influence which representation drives the observed effects (Cowan et al., 2013; Vergauwe & Cowan, 2015). The present results demonstrate that the type of stimulus (relational vs. conjunctive) and participants’ age are further variables that need to be considered in models of information integration in WM.

Given the results from Ecker et al. (2013), the virtually equivalent costs for task-irrelevant relational and conjunctive feature changes in younger adults were surprising. We can only speculate about the reasons for this difference. One relevant factor might lie in the relational stimuli’s composition. The items in the study of Ecker and colleagues were three-dimensional shapes that were superimposed on squares, with color manipulated as a foreground feature of the shape or a feature of the background square. In order to perceive the shape, the figure must be separated from the ground (see Peterson & Kimchi, 2013, for a
review). If attention is oriented to the foreground it cannot be oriented at the same time to the background. This would make it less likely to encode the irrelevant relational color together with the shape. Potentially, this hindered processing of background color and consequentially impaired memory for it.

Furthermore, due to the numerous potential moderators of the binding effect, age-related influences on binding should be investigated with other paradigms and materials. For example, an obvious difference to previous studies is that our encoding times were individually estimated (see Appendix). Therefore, the presentation times of the study display were longer than usual (e.g., Gao et al., 2011; Luck & Vogel, 1997; Xu, 2002b). Our results were similar to those of Peterson and Naveh-Benjamin (2016), who used shorter encoding times, and those of van Geldorp et al. (2015), who used longer encoding times. The same independence of encoding time was found for conjunctive stimuli (Brown et al., 2017; Rhodes et al., 2016). Thus, encoding time does not seem critical for the absence of the age-related binding deficit in the direct test of our study. However, if shape-color binding is a byproduct of feature encoding with both types of stimuli, encoding time should not influence performance provided it is sufficiently long to allow feature encoding, as no additional time-consuming binding processes are performed. This may be different if other types of information have to be associated which require additional binding processes beyond simultaneous feature encoding. If these processes are time consuming, binding should only be possible with long but not short encoding times (e.g., Rhodes et al., 2016) and the effect of encoding time may even become stronger with advancing age (Rhodes et al., 2017).

Finally, an open issue is how the test procedure might influence participants’ cognitive processing. We used a single test probe to focus on object-feature bindings and eliminate the influence of object-location bindings (Cowan et al., 2013). Participants were asked to decide whether the single test item matched their memory representation. It is
unknown whether this similarity-driven comparison is identical to processing in a whole-report change detection task. It is assumed that in the whole-report task, participants initially perform a parallel search for differences and subsequently engage in a conditional evaluation of the critical, potentially deviant item (Hyun, Woodman, Vogel, Hollingworth, & Luck, 2009). This secondary evaluation process appears similar to the evaluation of a single test probe. However, the evaluation of a single probe may have a stronger focus on detecting “sameness” (i.e., a match with the memory representation), whereas the secondary evaluation in a whole-report task may have a stronger focus on detecting a difference (i.e., a mismatch). We are not aware of any study that has tested this assumption directly; this thus remains a target for future research.
AGE-RELATED CHANGES TO RELATIONAL BINDING

References


https://doi.org/10.1016/j.neubiorev.2010.01.014


Table 1  
Recognition performance scores (Pr) for older and younger participants in the direct association WM test

<table>
<thead>
<tr>
<th>Age group</th>
<th>Stimulus type</th>
<th>Change type</th>
<th>Pr (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>Relation</td>
<td>New Shape</td>
<td>0.56 (.12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>New Color</td>
<td>0.62 (.14)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recombined</td>
<td>0.39 (.19)</td>
</tr>
<tr>
<td></td>
<td>Conjunction</td>
<td>New Shape</td>
<td>0.68 (.12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>New Color</td>
<td>0.71 (.17)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recombined</td>
<td>0.56 (.16)</td>
</tr>
<tr>
<td>Old</td>
<td>Relation</td>
<td>New Shape</td>
<td>0.44 (.16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>New Color</td>
<td>0.52 (.19)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recombined</td>
<td>0.29 (.19)</td>
</tr>
<tr>
<td></td>
<td>Conjunction</td>
<td>New Shape</td>
<td>0.52 (.15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>New Color</td>
<td>0.66 (.15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recombined</td>
<td>0.41 (.18)</td>
</tr>
</tbody>
</table>

*Note. Pr = Mean corrected recognition scores*
Table 2
Recognition performance scores (Pr) and response time (RT) costs for older and younger participants in the indirect irrelevant change WM test

<table>
<thead>
<tr>
<th>Age group</th>
<th>Stimulus type</th>
<th>Change type</th>
<th>Pr (SD)</th>
<th>RT costs (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>Relation</td>
<td>New Color</td>
<td>0.64</td>
<td>(0.17) 56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recombined</td>
<td>0.67</td>
<td>(0.15) 61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Same</td>
<td>0.75</td>
<td>(0.14)</td>
</tr>
<tr>
<td></td>
<td>Conjunction</td>
<td>New Color</td>
<td>0.64</td>
<td>(0.18) 76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recombined</td>
<td>0.63</td>
<td>(0.17) 53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Same</td>
<td>0.74</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Old</td>
<td>Relation</td>
<td>New Color</td>
<td>0.55</td>
<td>(0.17) 29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recombined</td>
<td>0.57</td>
<td>(0.14) 43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Same</td>
<td>0.60</td>
<td>(0.17)</td>
</tr>
<tr>
<td></td>
<td>Conjunction</td>
<td>New Color</td>
<td>0.44</td>
<td>(0.22) 160</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recombined</td>
<td>0.49</td>
<td>(0.17) 139</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Same</td>
<td>0.60</td>
<td>(0.15)</td>
</tr>
</tbody>
</table>

Note. Pr = Mean corrected recognition scores; RT costs = response time costs (in ms).
Figure 1. Schematic illustration of the trial structure, not drawn to scale. A “shape” cue was associated with the indirect test; a “both” cue was associated with the direct test. In the example, in the indirect test condition, a “no change” trial is depicted (i.e., the test probe is a studied shape; the color change is irrelevant and the correct response is thus “old”). For the direct test condition, a “change” trial is depicted (i.e., the test probe is a recombination of a studied shape and a studied color and the correct response is thus “new”). The width of the colored frames in the relational condition was determined by the surface area of the corresponding shape, in order to equate the number of color pixels across relational and conjunctive stimulus categories.
Figure 2. Corrected recognition scores ($Pr$) for older and younger participants in the direct test. Performance is depicted for item changes (aggregated across new color/new shape changes) and binding changes (recombination) across stimulus types (relation, conjunction). Error bars show 95% confidence intervals of means based on ipsative residuals (following Cousineau, 2005).
Figure 3. Corrected recognition scores ($Pr$) for older and younger participants in the direct test. Performance is depicted for item changes (new color, new shape) across stimulus types (relation, conjunction). Error bars show 95% confidence intervals of means based on ipsative residuals (following Cousineau, 2005).
Figure 4. Corrected recognitions scores ($Pr$) for older and younger participants in the indirect test. Performance is depicted for the irrelevant change conditions (new color, recombined) and the no change condition across stimulus types (relation, conjunction). Error bars show 95% confidence intervals of means based on ipsative residuals (following Cousineau, 2005).
Figure 5. Recognition performance costs of irrelevant relational and conjunctive feature changes for older and younger adults. Error bars show 95% confidence intervals of means based on ipsative residuals (following Cousineau, 2005).
Appendix

Visual search task

The design of the search task was a fully crossed 2 (stimulus type: relation, conjunction) × 2 (set size: 4, 9) design, with the target being present in 50% of the trials. The task had eight blocks with 10 trials each. Conditions appeared in random order.

Materials and Apparatus. The same shape stimuli as used in the main experiment were presented in the search task, except that shapes were always presented in white color in both relational and conjunctive format. Proper encoding was thus estimated according to the more complex shape information (Alvarez & Cavanagh, 2004).

Procedure. Participants were required to indicate whether a target shape was present or absent from a display of four or nine stimuli. Each trial began with the presentation of a fixation cross for 400 ms, which changed to an “x” for an interval of between 650 and 950 ms (exact duration was randomly determined) and then changed back to a fixation cross for another 400 ms. Thereafter, the target stimulus was presented for 500 ms at the center of the screen; then, after a blank interval of 900 ms, the search array of 4 or 9 stimuli was presented and remained on-screen until a response was given or for a maximum of 5,000 ms. Stimuli of the search array were randomly positioned in an invisible 3 × 3 grid. Participants made a present/absent response by pressing a key on a Cedrus response pad (RB-834, Cedrus Corporation). Response categories were assigned according to the handedness of the participant, with target-present responses mapped onto the dominant hand. Both accuracy and response speed were emphasized.

Encoding time per item (ETI) was estimated by calculating the mean RT for correct trials separately for each search array size, and then calculating a measure of search slope: 

$$ETI = \frac{(mean \ RT_{9 \ items} - mean \ RT_{4 \ items})}{5}$$

(see Alvarez & Cavanagh, 2004). Presentation time for the three study items of the main experiment was then set to $3 \times ETI + action \ onset \ time$. 
The action onset time was a constant time provided for action initiation and task set implementation. It was calculated as the mean RT for nine items minus the estimated search time for nine items, i.e. $\text{mean RT}_{9 \text{ items}} - (ETI \times 9)$. To ensure adequate understanding of the paradigm, all participants performed eight practice trials that were not included in later data analyses.

**Results.** Trials were excluded prior to analyses according to the same scheme applied to the main experiment, resulting in the exclusion of 0.47% of the trial data. Descriptive statistics for accuracy and RTs are given in Table A1.

Overall, participants solved the search task with high accuracy. Older adults had slightly lower performance, $M = .91$, 95% CI [.90, .92], $SE = .006$, than younger adults, $M = .94$, 95% CI [.93, .96], $SE = .006$, $U = 519.50$, $z = -3.75$, $p < .001$, but the differences between conditions were minor.

The analysis of search times revealed a different picture. RT data were analyzed in a $2 \times 2 \times 2 \times 2$ mixed-measures ANOVA with the within-subjects factors set size (4, 9), stimulus type (relation, conjunction), and target type (absent, present), and the between-subjects factor age (young, old). No main effect of set size was observed, $F < 1$. Search times were longer for relational stimuli, $M = 2,071$, 95% CI [1,911, 2,150], $SE = 40$, than for conjunctive stimuli, $M = 1,294$ ms, 95% CI [1,244, 1,343], $SE = 25$, $F(1, 86) = 1,248.91$, $p < .001$, $\eta_p^2 = .94$. Target-absent trials, $M = 1,976$ ms; 95% CI [1,898, 2,054], $SE = 39$, had longer RTs than target-present trials, $M = 1,389$ ms, 95% CI [1,333, 1,445], $SE = 28$, $F(1, 86) = 479.56$, $p < .001$, $\eta_p^2 = .85$. Older adults, $M = 1,882$ ms, 95% CI [1,792, 1,971], $SE = 45$, had longer search times than younger adults, $M = 1,483$ ms, 95% CI [1,395, 1,571], $SE = 44$, $F(1, 86) = 39.79$, $p < .001$, $\eta_p^2 = .32$. Moreover, there was a significant interaction of stimulus type and age, $F(1, 86) = 10.51$, $p = .002$, $\eta_p^2 = .11$: the RT difference between conjunctive and relational stimuli was larger for older adults, $M = 848$ ms, 95% CI Difference [786, 911], $SE$
AGE-RELATED CHANGES TO RELATIONAL BINDING

= 31, than for younger adults, $M = 706$ ms, 95% CI_Difference [644, 767], $SE = 31$. The interaction of target type and age was also significant, $F(1, 86) = 19.43$, $p < .001$, $\eta^2_p = .18$, indicating that the difference between target-present and target-absent trials was greater for older adults, $M = 705$ ms, 95% CI_Difference [629, 781], $SE = 38$, than for younger adults, $M = 469$ ms, 95% CI_Difference [394, 543], $SE = 38$. Finally, stimulus type and target type interacted, $F(1, 86) = 248.63$, $p < .001$, $\eta^2_p = .74$. The RT difference between conjunctive and relational stimuli was larger for target-absent trials, $M = 1,082$ ms, 95% CI_Difference [1,016, 1,147], $SE = 33$, than for target-present trials, $M = 472$ ms, 95% CI_Difference [422, 521], $SE = 25$.

**Discussion.** The data were as expected. Participants across both age groups showed high levels of performance. However, older participants had longer search times, and this effect was more pronounced if the perceptual task was difficult. Thus, older participants were given longer encoding times, adapted to their personal search time, in order to compensate for perceptual slowing.
Table A1
Recognition performance and response times for younger and older adults in the search task

<table>
<thead>
<tr>
<th>Set size</th>
<th>Stimulus</th>
<th>Target</th>
<th>Younger ACC</th>
<th>Younger RT</th>
<th>Older ACC</th>
<th>Older RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Relation</td>
<td>absent</td>
<td>.98 (.05)</td>
<td>2,202 (464)</td>
<td>.88 (.14)</td>
<td>2,839 (526)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>present</td>
<td>.90 (.10)</td>
<td>1,466 (366)</td>
<td>.89 (.10)</td>
<td>1,770 (346)</td>
</tr>
<tr>
<td></td>
<td>Conjunction</td>
<td>absent</td>
<td>.98 (.04)</td>
<td>1,242 (277)</td>
<td>.93 (.08)</td>
<td>1,642 (287)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>present</td>
<td>.93 (.08)</td>
<td>1,055 (243)</td>
<td>.94 (.08)</td>
<td>1,277 (247)</td>
</tr>
<tr>
<td>9</td>
<td>Relation</td>
<td>absent</td>
<td>.97 (.06)</td>
<td>2,210 (489)</td>
<td>.88 (.13)</td>
<td>2,815 (567)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>present</td>
<td>.90 (.11)</td>
<td>1,465 (421)</td>
<td>.88 (.11)</td>
<td>1,798 (359)</td>
</tr>
<tr>
<td></td>
<td>Conjunction</td>
<td>absent</td>
<td>.97 (.05)</td>
<td>1,216 (262)</td>
<td>.94 (.07)</td>
<td>1,640 (283)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>present</td>
<td>.93 (.09)</td>
<td>1,009 (231)</td>
<td>.93 (.10)</td>
<td>1,272 (272)</td>
</tr>
</tbody>
</table>

Note. ACC = Accuracy; RT = response time in ms.