

Replicating distinctive facial features in lineups: identification performance in young versus older adults

Stephen P. Badham · Kimberley A. Wade ·
Hannah J. E. Watts · Natalie G. Woods ·
Elizabeth A. Maylor

Published online: 13 November 2012
© Psychonomic Society, Inc. 2012

Abstract Criminal suspects with distinctive facial features, such as tattoos or bruising, may stand out in a police lineup. To prevent suspects from being unfairly identified on the basis of their distinctive feature, the police often manipulate lineup images to ensure that all of the members appear similar. Recent research shows that replicating a distinctive feature across lineup members enhances eyewitness identification performance, relative to removing that feature on the target. In line with this finding, the present study demonstrated that with young adults ($n = 60$; mean age = 20), replication resulted in more target identifications than did removal in target-present lineups and that replication did not impair performance, relative to removal, in target-absent lineups. Older adults ($n = 90$; mean age = 74) performed significantly worse than young adults, identifying fewer targets and more foils; moreover, older adults showed a minimal benefit from replication over removal. This pattern is consistent with the associative deficit hypothesis of aging, such that older adults form weaker links between faces and their distinctive features. Although replication did not produce much benefit over removal for older adults, it was not detrimental to their performance. Therefore, the results suggest that replication may not be as beneficial to older adults as it is to young adults and demonstrate a new practical implication of age-related associative deficits in memory.

Keywords Police lineups · Face identification · Distinctive features · Aging · Associative deficits

Individuals with distinctive facial markings (e.g., moustache, tattoo) often stand out from a crowd. When creating lineups for criminal investigations, police must ensure that suspects with distinctive features do not stand out among other lineup members (or *foils*). Indeed, erroneous identifications are more likely to occur when the suspect—who may or may not be the actual culprit—is the only member of the lineup who matches the eyewitness's description of the culprit (Doob & Kirshenbaum, 1973; Wells, Rydell, & Seelau, 1993). This is especially true for simultaneous lineups, in which lineup members are presented all together, and witnesses typically select the member who looks *most* like the culprit they hold in memory, even if that member is not the same person (Wells et al., 1998).

The problem of suspects standing out in lineups is surprisingly common; around one third of all lineups in England and Wales are digitally manipulated to prevent suspects from standing out (see Zarkadi, Wade, & Stewart, 2009), and in a survey of U.S. police officers, 70 % reported using methods to avoid a suspect with distinctive features from standing out (Wogalter, Malpass, & McQuiston, 2004). Image manipulation is a sensible and practical solution for creating lineups for suspects with distinctive features: It is relatively quick and inexpensive, and the police can create multiple foils with idiosyncratic features, such as tattoos, piercings, or bruising. Yet there are no strict guidelines in the U.K. or in the U.S. on how the police should manipulate images; the technique used is typically left to the investigating officer to decide, and this judgment is frequently based on personal experience and intuition (Zarkadi et al., 2009).¹

S. P. Badham (✉) · K. A. Wade · H. J. E. Watts · N. G. Woods ·
E. A. Maylor
Department of Psychology, University of Warwick,
Coventry CV4 7AL, UK
e-mail: s.p.badham@warwick.ac.uk

S. P. Badham
College of Medicine, Biological Sciences and Psychology,
Henry Wellcome Building, University of Leicester,
Lancaster Road,
Leicester LE1 9HN, UK

¹ All of the facial recognition/lineup studies cited in the introduction used photographic stimuli, which is in line with three quarters of real police lineups (Wogalter et al., 2004).

Zarkadi and colleagues (2009) recently compared two methods for preventing suspects from standing out—namely, *replication*, in which the target's distinctive feature is digitally added to all of the foils, and *removal*, in which the target's distinctive feature is removed and the target appears among foils without distinctive features. Testing young (college-aged) adults, Zarkadi et al. (2009) found that replication resulted in more target identifications than did removal in target-present (TP) lineups, without increasing foil identifications in target-absent (TA) lineups. This finding fits with the encoding specificity hypothesis (Tulving & Thompson, 1973), which states that memory performance is improved when encoding and retrieval occur in similar contexts. Presumably, replication enhances identification performance because the target appears exactly as it did during encoding, but for removal it does not. This enhancement on TP lineups is also predicted by Valentine's (1991) face-space framework and the hybrid-similarity model of recognition (Nosofsky & Zaki, 2003).

However, age-related differences in memory may influence how young and older adults remember faces with distinctive features, and this may impact on the most suitable method of presenting lineups to older adults. Older adults (60+ years of age) generally have poorer episodic memory than do young adults (e.g., Zacks, Hasher, & Li, 2000), and this applies to memory for faces (e.g., Bartlett, Leslie, Tubbs, & Fulton, 1989; Grady et al., 1995; Naveh-Benjamin, Guez, Kilb, & Reedy, 2004). Using an eyewitness memory task, Memon, Hope, Bartlett, and Bull (2002) found that older adults were more prone to selecting foils in TA lineups than were young adults. They argued that older adults were more likely to respond on the basis of familiarity than were young adults. Similarly, in an old/new recognition memory test, Bartlett et al. found that older adults performed worse than young adults because they identified more new faces at test. These results indicate that older adults are more prone to making erroneous positive facial identifications than are young adults. Therefore, unlike young adults, they may be more likely to identify foils in replication lineups, as compared with removal lineups, because replication foils are more familiar since they share distinctive features with the target. Thus, replication may be detrimental to older adults' memory, as compared with removal.

In addition, a well-established theory of cognitive aging suggests that replication is less likely to benefit older adults, as compared with removal, but that replication will not hinder their performance. Older adults show deficits in forming associations between units of information (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008). As a result, older adults show smaller deficits in memory tests for

units of information, such as recognizing old versus new words, than for associations between units of information, such as recognizing intact versus rearranged word pairs. Age-related associative deficits have been demonstrated in studies requiring the binding of faces to temporal contexts (Bastin & Van der Linden, 2005), the binding of two simultaneously presented faces and the binding of faces to spatial positions (Bastin & Van der Linden, 2006), and the binding of faces to names (Naveh-Benjamin et al., 2004; Naveh-Benjamin et al., 2009). If older adults express associative deficits when forming links between faces and their distinctive features, the benefit of replication may be reduced. This is because replication is likely to rely on reinstating the links between faces and their distinctive features at test to cue memory for those links; thus, if the memory is not present, such a cue will not be effective.

The present study examined whether replication is the better technique for creating lineups for both young and older eyewitnesses. Zarkadi et al.'s (2009) laboratory-based recognition memory task was adapted to test older as well as young adults. If increased reliance on familiarity dominates age differences in this type of memory task, we would expect to see replication as detrimental to older adults' performance, as compared with removal. However, if age-related associative memory deficits are the primary influence on performance, older adults should show similar memory performance with replication and removal or, at least, a smaller benefit from replication over removal than young adults. On a practical level, this research could determine whether replication (in preference to removal) should be recommended to police officers conducting lineups for adult witnesses of all ages.

Method

Participants

Sixty young adults (30 female) 18–24 years of age ($M = 20.4$, $SD = 1.4$) and 90 older adults (51 female) 61–91 years of age ($M = 74.2$, $SD = 7.4$) participated voluntarily. Young participants were an opportunity sample. Older participants were recruited from the University of Warwick Age Study volunteer panel and from the local community. We used a 2 (age: young, older) \times 2 (lineup type: replication, removal) \times 2 (target presence: present, absent) mixed design, with lineup type and target presence manipulated within participants.

Participants completed two measures of cognitive functioning that are standard in the aging literature: the digit symbol substitution task (Wechsler, 1981), a

measure of perceptual–motor processing speed, and the multiple-choice section of the Mill Hill vocabulary test (Raven, Raven, & Court, 1988), a measure of crystallized intelligence. The results were consistent with those in the literature (e.g., Horn & Cattell, 1967; Salthouse, 1991). Young adults were significantly faster than older adults on the speed task, $t(148) = 12.21, p < .001$ ($M_{\text{young}} = 67.0, SD = 12.5$; $M_{\text{older}} = 42.5, SD = 11.6$), and scored significantly lower than older adults on the vocabulary test, $t(148) = 3.79, p < .001$ ($M_{\text{young}} = 16.5, SD = 4.4$; $M_{\text{older}} = 19.7, SD = 5.5$).

Stimuli

The stimuli were grayscale images of 98 faces used by Zarkadi et al. (2009). The images were from Florida’s Department of Corrections Web site and were of inmates 24 years of age with short brown hair, brown eyes, and neutral expressions. Inmates were looking directly at the camera and were in front of a uniform gray background. Zarkadi et al. (2009) used Adobe Photoshop CS2 to digitally remove any distinguishing features, such as birthmarks or facial hair. They then selected 42 of the 98 faces, split them into six subsets of 7 faces, and randomly assigned each subset a distinctive feature (bruise, mole, moustache, piercing, scar, and tattoo) to be digitally added to the face (as illustrated in Fig. 1). Note that in the present study, the distinctive features were discrete markings applied to otherwise normal faces, not unusual facial features such as a large nose.

The memory set consisted of 32 faces: 26 preselected, nondistinctive faces and 6 randomly selected target faces that had one of each of the six distinctive features. The remaining stimuli were used to randomly create 12 six-person lineups. There were 6 TP and 6 TA lineups crossed with replication and removal to complete the 2×2 design. Hence, 3 lineups were created for each of the four lineup conditions.

Replication TP lineups

A target (distinctive) face from study appeared in the lineup with a distinctive mark, and a similar mark was replicated across all of the foils.

Removal TP lineups

A target (distinctive) face from study appeared in the lineup without a distinctive mark, and faces with no distinctive marks were used as foils.

Replication TA lineups

No target face from study appeared in the lineup. All six lineup members were new (unseen) foils with a distinctive feature that was similar to one of the six distinctive features presented at study.

Removal TA lineups

No target face from study appeared in the lineup. All six lineup members were new (unseen) foils without distinctive features.

Procedure

At study, participants were told that they would view 32 faces, 1 at a time for 3 s each, and that their memory for these faces would be tested. Participants were instructed to remember the individuals themselves because they might appear differently in the memory test. Of the 32 faces, the 6 faces with distinctive marks appeared in the test phase as targets, and the remaining 26 faces without distinctive features were not seen again. The 32 faces were presented in random order. After the study phase, participants completed the digit symbol substitution task for a fixed duration of 90 s.

Fig. 1 Examples of distinctive features (*top*) digitally added to faces (*bottom*). From left to right: bruise, mole, moustache, piercing, scar, tattoo



At test, participants completed a lineup identification task. Participants were told that they would view 12 six-person lineups and their task was to indicate, via a buttonpress of numbers one to six on the laptop keyboard, which face they had seen before or, if they recognized no faces, to press the number zero. They were informed that they could respond only once and that there would not always be a face from the memory set in the lineup. The 12 lineups were presented in random order for each participant. There was no time limit for participants' decisions, and no feedback was provided.

Results and discussion

Participants' responses were categorized into three groups: *Target* identifications occurred when participants correctly identified a target face in a TP lineup, *foil* identifications occurred when they incorrectly identified foils in TP or TA lineups, and *none* responses occurred when they correctly (TA lineups) or incorrectly (TP lineups) decided that none of the faces had been seen before. Figure 2 shows the proportion of responses falling into each category for the four types of lineup and for young and older adults. Responses for each category (*target*, *foil*, and *none*) were entered individually into 2 (age: young, older) \times 2 (lineup type: replication, removal) repeated measures ANOVAs separately for TP and TA lineups.

TP lineups

More targets were identified in the replication lineups than in the removal lineups, $F(1, 148) = 32.85$, $MSE = 0.07$, $p < .001$, replicating the findings from Zarkadi et al. (2009). Target identifications were higher in young adults than in older adults, $F(1, 148) = 19.95$, $MSE = 0.11$, $p < .001$. Importantly, there was an interaction between age and lineup type, $F(1, 148) = 8.74$, $MSE = 0.07$, $p < .01$, with older adults benefiting less from the replication technique, as compared with young adults.

Foil identifications and none responses are both errors in TP lineups. More foils were identified under removal than under replication, $F(1, 148) = 15.81$, $MSE = 0.06$, $p < .001$. Foil identifications were lower in young adults than in older adults, $F(1, 148) = 16.59$, $MSE = 0.12$, $p < .001$, which fits with previous facial identification research (e.g., Memon et al., 2002). Given that the correct face was available as an option, these results suggest that older adults may be less able to apply a relative judgment strategy. There was an interaction between age and lineup type, $F(1, 148) = 6.06$, $MSE = 0.06$, $p < .05$, again with older adults benefiting less from replication, as compared with young adults. More *none* selections were made in the removal than in the replication lineups,

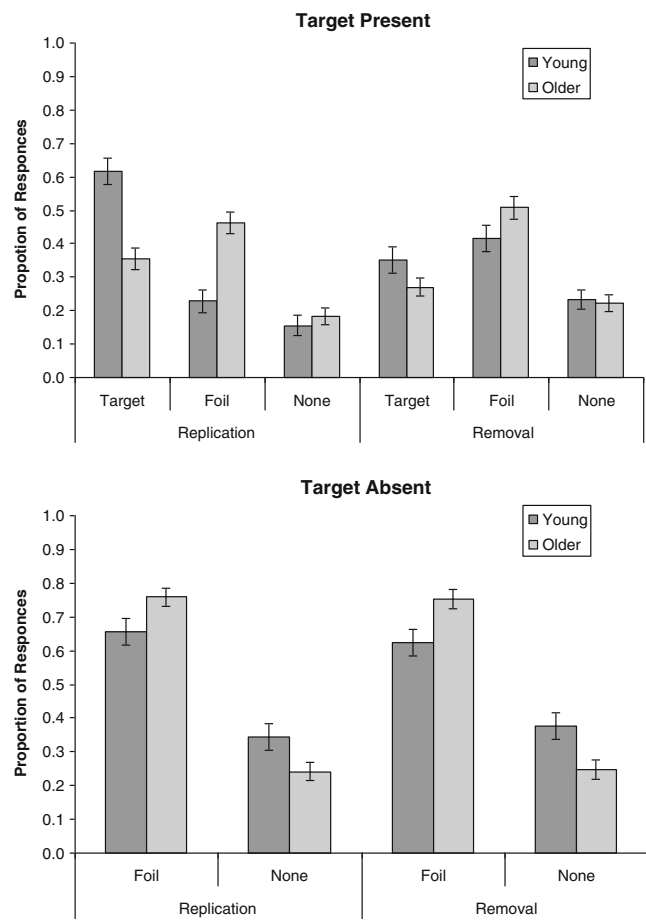


Fig. 2 Mean proportion of responses in each response category (identifying a target, a foil, or none of the faces) in replication and removal lineups for young and older adults and for target-present (*top*) and target-absent (*bottom*) trials. Error bars are ± 1 SE

$F(1, 148) = 5.33$, $MSE = 0.05$, $p < .05$, again indicating that replication lineups result in better performance. None responses were similar in young and older adults, $F < 1$, and there was no interaction between age and lineup type, $F < 1$.

To further establish whether young and older adults benefited from replication, we conducted a series of *t*-tests within each age group for TP lineups. Young adults performed better for all response types in replication than in removal lineups, identifying more targets, $t(59) = 5.27$, $p < .001$, and fewer foils, $t(59) = -4.19$, $p < .001$, and making fewer none responses, $t(59) = -2.12$, $p < .05$. Older adults, however, showed better performance for replication on target identifications, $t(89) = 2.30$, $p < .05$, but not for foil identifications, $t(89) = -1.19$, n.s., or none responses, $t(89) = -1.20$, n.s.

TA lineups

Foil identifications (errors) were lower in young adults than in older adults, $F(1, 148) = 9.43$, $MSE = 0.10$, $p < .01$. There was

no main effect of lineup type and no interaction between age and lineup type, $F_s < 1$. *None* responses are correct responses in TA lineups and were statistically identical, since *foil* and *none* response proportions must sum to one.

These results are consistent with the age-related associative deficit hypothesis, which predicts that older adults will form weaker links between faces and distinctive features. Forming an association between a face and a distinctive feature should have minimal impact on removal lineups because that distinctive feature is not present as a cue. Therefore, age-related associative deficits are more likely to occur in replication lineups where associative memories are important. Our data fit this prediction: For TP lineups, although there was superior performance overall for replication lineups than for removal lineups, older adults benefited to a significantly lesser extent than did young adults. Moreover, older adults did not show worse performance for replication than for removal lineups, despite the fact that foils have more familiar features in replication lineups. This indicates that age differences in reliance on familiarity are unlikely to have driven the present pattern of results.

One counterexplanation for our findings is that older adults benefitted less from replication because they found the task difficult and performed closer to chance levels. Indeed, there is evidence for an own-age bias in facial recognition, where people are not as good at recognizing faces of different ages from themselves (e.g., Perfect & Moon, 2005; Wright & Stroud, 2002). Our older adults may have struggled more with the task because they were required to identify a series of young (24 years of age) faces. To address this issue more closely, we examined the proportion of correct positive responses in TP lineups (i.e., $target/[target + foil]$ identifications) across the two age groups (see Fig. 3 for means).² The pattern of results was similar to that of our main analysis. Young adults identified a larger proportion of targets than did older adults, $F(1, 146) = 17.65$, $MSE = 0.15$, $p < .001$, and there was an age \times lineup type interaction, $F(1, 146) = 7.26$, $MSE = 0.10$, $p < .01$. Paired t -tests between replication and removal lineups confirmed that young adults benefited from replication, $t(59) = 4.40$, $p < .001$, but unlike in the earlier analysis including none responses, older adults did not, $t(87) = 1.31$, n.s. Moreover, our data show that the older adults performed significantly better than chance. Given that a selection was made, the chance of identifying a target was 1/6, or .17. The proportion of target endorsements by older adults significantly exceeded this level on both replication, $t(88) = 7.07$,

² Two older adults made no endorsements for one of the lineup types with TP lineups (one for replication lineups and one for removal lineups), so were not included in this analysis.

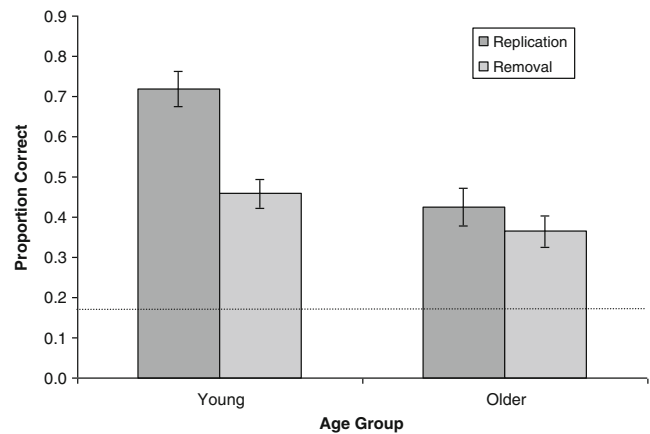


Fig. 3 Proportion of endorsements in target-present lineups that were correct for replication and removal lineups for young and older adults. The dashed line indicates chance performance; error bars are ± 1 SE

$p < .001$, and removal, $t(88) = 5.25$, $p < .001$, lineups.³ It is also worth noting that young and older adults did not differ significantly (all $p_s > .05$) on removal lineups in terms of either correct target endorsements (Fig. 3) or target, foil, and none responses (Fig. 2), demonstrating that the older adults did not find parts of the task significantly more difficult than did the young adults. Together, these findings suggest that the age differences observed are not due to older adults performing at chance level or finding the task overly difficult.⁴

As a further check that the difference in benefit for replication over removal between young and older adults was not due to overall memory performance, the difference in the proportion of target endorsements between replication and removal lineups was calculated and compared with two independent measures of memory performance: (1) the proportion of none responses in all TA lineups and (2) the proportion of none responses in all TP lineups. Neither of the two measures correlated with the difference in performance between replication and removal lineups, $r(148) = -.002$, $p = .98$, and $r(148) = -.050$, $p = .54$, respectively. Therefore, the benefit of replication lineups over removal lineups (as measured by

³ A further analysis was conducted after removing the 31 % of older adults who performed at or below chance, as well as the same percentage of lowest performing young adults. The results produced identical patterns of significance to those reported above.

⁴ Note also that Zarkadi et al.'s (2009) original task was more difficult than the present version, due to a shorter presentation time and a longer delay between study and test. In terms of proportions of responses to targets in the removal condition, the present older adults and Zarkadi et al.'s (2009) young adults were almost perfectly matched (.27 and .28, respectively), whereas in the replication condition, the present older adults performed worse than Zarkadi et al.'s (2009) young adults (.36 and .49, respectively).

target identifications) is not purely determined by overall memory performance.

The overall data pattern was consistent with that in Zarkadi et al. (2009): Replication resulted in superior target identification, as compared with removal, without leading to a corresponding increase in foil identification in TA lineups. In the present study, for young adults, the overall level of performance was higher than in Zarkadi et al.'s (2009) study for TP lineups. This result is unsurprising, since study faces were presented for 3 s each in the present study, rather than 2 s each as in Zarkadi et al.'s (2009) study, and the delay between study and test was 90 s in the present study but was 5 min in Zarkadi et al.'s (2009) study. For TA lineups, the level of performance was similar across the two studies, which is less easy to explain since the present study should have shown improved performance, as compared with Zarkadi et al.'s (2009) study.

Interestingly, replication did not enhance identification performance in TA lineups. However, Zarkadi et al. (2009) and Zarkadi, Stewart, and Wade (2012) also found no effect of lineup technique in TA lineups with young adults. Recently, these authors used their TA data to test predictions made by recognition models based on global familiarity. Modeling revealed that identification in lineups for suspects with distinctive features is based on a one-to-one match to the memory of the perpetrator, and not on global familiarity to other faces. This finding may explain why lineup technique does not affect performance on TA lineups and presents a challenge to some global familiarity models of face recognition (e.g., the face-space model of Valentine, 1991).

The present results were from a laboratory-based face recognition experiment that was difficult for all participants. Now that differences in replication and removal lineups have been established, future research should aim to extend these findings in a more ecologically valid eyewitness identification paradigm. Such a design may involve viewing a video of a crime rather than a sequence of face images, then viewing just one lineup after a long delay. Additionally, to further investigate the role of age-related associative deficits on the effects of replication and removal, future work should examine whether the same pattern of age differences occurs under incidental study.⁵ It has been shown that the age-related associative deficit is reduced under incidental encoding conditions (Old & Naveh-Benjamin, 2008). Therefore, an incidental learning paradigm may cause older adults to benefit more from replication over removal; or perhaps the inverse may occur, and the benefit of replication might be reduced in young adults.

To conclude, the present results indicate that age-related associative deficits not only have an impact on overall memory performance, but also can influence the qualitative pattern of older adults' behavior. In terms of eyewitness identification,

researchers and practitioners should not assume that a procedure that is beneficial to young adults (in this case, replication rather than removal in the creation of police lineups) will necessarily be as effective in older adults.

Author Note This research formed part of a doctoral dissertation by Stephen P. Badham, funded by a University of Warwick Postgraduate Research Fellowship. We are grateful to Theodora Zarkadi for providing the stimuli and to Neil Stewart for comments on an earlier version of the manuscript.

References

- Bartlett, J. C., Leslie, J. E., Tubbs, A., & Fulton, A. (1989). Aging and memory for pictures of faces. *Psychology and Aging, 4*, 276–283.
- Bastin, C., & Van der Linden, M. (2005). Memory for temporal context: Effects of ageing, encoding instructions, and retrieval strategies. *Memory, 13*, 95–109.
- Bastin, C., & Van der Linden, M. (2006). The effects of aging on the recognition of different types of associations. *Experimental Aging Research, 32*, 61–77.
- Doob, A. N., & Kirshenbaum, H. (1973). Bias in police lineups—partial remembering. *Journal of Police Science and Administration, 1*, 287–293.
- Grady, C. L., McIntosh, A. R., Horwitz, B., Maisog, J. M., Ungerleider, L. G., Mentis, M. J., et al. (1995). Age-related reductions in human recognition memory due to impaired encoding. *Science, 269*, 218–221.
- Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*, 107–129.
- Memon, A., Hope, L., Bartlett, J., & Bull, R. (2002). Eyewitness recognition errors: The effects of mugshot viewing and choosing in young and old adults. *Memory & Cognition, 30*, 1219–1227.
- Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1170–1187.
- Naveh-Benjamin, M., Guez, J., Kilb, A., & Reedy, S. (2004). The associative memory deficit of older adults: Further support using face–name associations. *Psychology and Aging, 19*, 541–546.
- Naveh-Benjamin, M., Shing, Y. L., Kilb, A., Werkle-Bergner, M., Lindenberger, U., & Li, S.-C. (2009). Adult age differences in memory for name–face associations: The effects of intentional and incidental learning. *Memory, 17*, 220–232.
- Nosofsky, R. M., & Zaki, S. R. (2003). A hybrid-similarity exemplar model for predicting distinctiveness effects in perceptual old-new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1194–1209.
- Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23*, 104–118.
- Perfect, T. J., & Moon, H. C. (2005). The own-age effect in face recognition. In J. Duncan, P. McLeod, & L. H. Phillips (Eds.), *Measuring the mind: Speed, control, and age* (pp. 317–340). New York: Oxford University Press.
- Raven, J. C., Raven, J., & Court, J. H. (1988). *The Mill Hill vocabulary scale*. London: H. K. Lewis.
- Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale: Lawrence Erlbaum.
- Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.
- Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of*

⁵ We thank one of the reviewers for this suggestion.

- Experimental Psychology. A, Human Experimental Psychology*, 43, 161–204.
- Wechsler, D. (1981). *Manual for the Wechsler adult intelligence scale—revised*. New York: Psychological Corporation.
- Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distracters for eyewitness lineups. *Journal of Applied Psychology*, 78, 835–844.
- Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behaviour*, 22, 603–647.
- Wogalter, M. S., Malpass, R. S., & McQuiston, D. E. (2004). A national survey of us police on preparation and conduct of identification lineups. *Psychology, Crime and Law*, 10, 69–82.
- Wright, D. B., & Stroud, J. N. (2002). Age differences in lineup identification accuracy: People are better with their own age. *Law and Human Behaviour*, 26, 641–654.
- Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In T. A. Salthouse & F. I. M. Craik (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 293–357). Mahwah: Lawrence Erlbaum Associates, Inc.
- Zarkadi, T., Wade, K. A., & Stewart, N. (2009). Creating fair lineups for suspects with distinctive features. *Psychological Science*, 20, 1448–1453.
- Zarkadi, T., Stewart, N., & Wade, K. A. (2012). *Lineups for suspects with distinguishing marks: Eyewitness identification is based on matching one face, not global familiarity*. Manuscript submitted for publication.