Memory & Cognition
2005, 33 (6), 1017-1024

Face recognition and aging: Effects of target age and memory load

ALLISON C. LAMONT
Massey University, Palmerston North, New Zealand

STEVE STEWART-WILLIAMS
McMaster University, Hamilton, Ontario, Canada

and

JOHN PODD
Massey University, Palmerston North, New Zealand

In the present study, we investigated age-related decline in face recognition memory and whether this decline is moderated by the age of the target faces and by the number of faces that the participant must learn (memory load). Thirty-two participants in each of three age groups (18–39 years, 60–75 years, and 76–96 years) completed a face recognition task. Signal detection analyses confirmed that face recognition accuracy declined with age. However, this finding was qualified by an interaction between participant age and target age, which revealed that the age-related decline in face recognition accuracy occurred only for young target faces. Increased memory load was associated with comparable performance decrements across all age groups. However, memory load appears not to be the cause of these decrements. Instead, they appear to be a product of recognition load (the number of stimuli presented in the recognition phase).

Humans learn to recognize thousands of faces during the course of their lives. The vast number of faces we learn, coupled with the high level of similarity of these stimuli, led Damasio (1989) to suggest that learning and recognizing faces is one of the most cognitively demanding and neurologically complex tasks in which we engage. Face recognition has been intensively investigated over the last few decades, resulting in a voluminous body of literature. A large amount of research has investigated the effects of participant age on face recognition. However, little attention has been given to variables that may interact with age in determining face recognition performance. In the present study, we examined two possible interacting variables: (1) the age of the target face and (2) *memory load* (the number of faces that the participants are required to learn). The main questions we investigated were: Do people of different ages have differential recognition success with target faces of different ages? Does increasing memory load have a greater impact on the elderly than on younger individuals, or does it have approximately the same impact across the life span? In addition, we were interested in determining whether the effects of memory load can be distinguished from those of *recognition load* (the total number of target and distractor faces seen in the recognition phase).

Adults can successfully encode large numbers of new faces from briefly inspected photographs and subsequently identify these from distractors at hit rates of over 90% (Carey, 1992; Goldstein, 1977). However, with increasing age come declines in various cognitive abilities, including face-recognition accuracy (Salthouse, 2004; Shapiro & Penrod, 1986). Researchers have shed light on various aspects of this cognitive decline, including its time course. Crook and Larrabee (1992), for instance, found significant age-related decrements in participants as young as 50, but found that the largest decrements occurred over the age of 70. This indicates that memory decline in face recognition is not linear but accelerates after the age of 70 and is consistent with evidence suggesting that memory in general deteriorates more rapidly in those over 70 years of age than in those a decade or so younger (Parkin, 1993). Researchers have also shed light on the nature of the deficits in face recognition with age. It appears that young adults and elderly people have similar hit rates, but that elderly people have an elevated level of false alarms (Bartlett & Leslie, 1986; Crook & Larrabee, 1992; Ferris, Crook, Clark, McCarthy, & Rae, 1980; Fulton & Bartlett, 1991; Smith & Winograd, 1978; but see Bäckman, 1991).

Although there is good evidence that increasing age is associated with a decline in the ability to recognize faces, little is known about the variables that might interact with

This research was funded by the School of Psychology at Massey University. All materials and procedures were approved by the Massey University Human Ethics Committee, Palmerston North (Protocol 01/28). Parts of this article were presented in 2003 at the 15th Annual Convention of the American Psychological Society in Atlanta. Correspondence should be addressed to J. Podd, School of Psychology, Massey University, Turitea Campus, Palmerston North, New Zealand (e-mail: j.v.podd@massey.ac.nz).

age. For instance, only a handful of studies have considered how the age of the participant might interact with the age of the target face. Nonetheless, there are theoretical reasons to expect such an interaction. Consider the recognition of same-race versus other-race faces. People generally find it easier to recognize faces of members of their own racial group than of members of other racial groups (Brigham & Barkowitz, 1978; Feinman & Entwisle, 1976; Lavrakas, Buri, & Mayzner, 1976; Shepherd, 1981; Shepherd, Deregowski, & Ellis, 1974). Associated with the cross-race effect is the often reported feeling that other-race faces "all look alike" (O'Toole, Peterson, & Deffenbacher, 1996). It is possible that an analogous effect occurs with the identification of people of the same age versus those of other ages. That is, young adults may be more accurate in identifying young faces than old faces, whereas older adults may be more accurate in identifying old faces than young faces.

There is some evidence supporting this proposition. Bäckman (1991) found that young adults ($M = 23.8$ years) showed better recognition for young faces than for old faces but that young-old adults ($M = 68.5$ years) showed better recognition for old faces than for young faces. The evidence pertaining to this issue is equivocal, however. Bäckman's study also included two groups of older participants (76 and 85 years), neither of which showed any effect of age of target face. Similarly, Bartlett and Fulton (1991) showed that recognition accuracy among young adults was generally better with young adult faces than with older faces, but that recognition accuracy among the elderly showed no significant effects of target age. Thus, the few studies that have investigated this issue suggest that young participants are more accurate in recognizing same-age faces, but they have found mixed evidence on the question of whether face age has any impact on elderly participants. Nonetheless, the cross-race effect in face identification provides some reason to expect that younger people will show better recognition of young than old faces and that older people will do the reverse. In the present study, we sought to add to the literature addressing this question.

Another relatively neglected variable in the face recognition literature is memory load (Goldstein & Chance, 1981; Metzger, 2002; Podd, 1990). There is some evidence that as memory load increases, face recognition performance decreases (or at any rate, false alarms increase; see Shapiro & Penrod, 1986). For instance, Podd had between 20 and 50 target faces in the exposure phase of a face recognition experiment. These values were chosen as representative of memory loads used in face recognition studies available at that time. Podd found reliable decreases in performance even with this relatively small increase in the number of target faces, indicating that memory load is an important factor to consider when interpreting the results of face recognition studies.

Podd (1990) did not consider participant age as a variable. A subsequent study by Metzger (2002) aimed to determine whether memory load and participant age interacted. Participants were presented with 10, 20, or 30 target faces during the exposure phase of the study and 20, 40, or 60 faces in the recognition phase. Three age groups were utilized: children ($M = 9.3$ years), college students ($M = 22$ years), and middle-aged adults ($M = 49.6$ years). Although analysis of the hit rate yielded statistically significant main effects for both memory load and participant age, no interaction between the two variables emerged. This provides some evidence that the effects of memory load are independent of age—that is, memory load does not have progressively larger effects in older participants. Given that Metzger's oldest group averaged only around 50 years of age, though, it is not certain that this result can be generalized to older age groups, especially in light of the increased rate of cognitive decline found among individuals older than 70 (Crook & Larrabee, 1992). Therefore, another aim of the present study was to investigate the effects of memory load in elderly participants.

A difficulty associated with research into memory load is that, when memory load increases, recognition load usually increases as well (Podd, 1990). Recognition load is the total number of target and distractor faces seen in the recognition phase. Typically, there are as many distractors as there are targets, and thus the recognition phase is twice as long as the exposure phase. Any effects associated with increased memory load (number of target faces in the exposure phase) might in fact be due to increased recognition load. Podd's study, the first to consider this potential confound, was designed to enable memory load to be assessed independently of the potentially confounding effects of recognition load. Participants were exposed to 20, 35, or 40 faces during the exposure face (40, 70, or 100 faces at recognition, respectively). To decouple memory load from recognition load, Podd ensured that the first 40 faces in the recognition phase (20 targets and 20 distractors) were identical and presented in the same order, for each group. By comparing between-group recognition rates for the first 40 recognition trials only, it was possible to assess the effects of memory load while holding recognition load constant. The results suggested that memory load, rather than recognition load, accounts for the decline in recognition performance. Metzger (2002), using a similar strategy, reached the same conclusion. Again, it is not known whether this finding applies to elderly (>75 years) participants. It is possible that an increased recognition load might be more problematic for elderly than for younger participants, due to the increased likelihood of interference and fatigue, factors more likely to affect older participants.

Hypotheses

On the basis of the research and theory outlined above, three hypotheses were formulated. Hypothesis 1 is that face recognition ability will decline with age, due largely to an increased rate of false alarms rather than a decreased hit rate. This hypothesis derives from previous research and constitutes a simple replication of past findings. Hypothesis 2 is that older people will be more accurate at recognizing older faces as opposed to younger faces, whereas younger people will be more accurate at recognizing younger faces as opposed to older faces. As discussed earlier, the evidence on this issue is equivocal. Nonetheless, the cross-race effect in face identification provides some reason to expect this result. Finally, Hypothesis 3 is that higher memory load will lead to performance decrements in face recognition accuracy and will produce greater deficits in elderly participants. This hypothesis is based on the supposition that older participants are more susceptible to interference and fatigue. To address these hypotheses, we varied participant age and memory load (between-groups factors) and target age (within-group factor) in a mixed factorial design.

METHOD

Participants

The participants were divided into three age groups: 18–39 years ($M = 25.93$, $SD = 5.85$), 60–75 years ($M = 66.84$, $SD = 4.69$), and 76 years and over ($M = 81.22$, $SD = 5.52$). There were 32 people in each group, 8 men and 24 women. The young participants were a mixture of college students and working people. The participants in the two older age groups were recruited from church and community groups; this selection process was similar to that used in previous research on face recognition and aging (e.g., Bartlett & Fulton, 1991). Half the participants within each age group were randomly assigned to the low memory load (LML) condition, and half were assigned to the high memory load (HML) condition. The randomization was provisional on maintaining the same gender balance across all groups; thus, there were 4 men and 12 women in each of the resulting six groups.

Task and Stimuli

As discussed earlier, one of the goals of the present study was to investigate the influence of memory load on face recognition accuracy. To deal with the potential confound of recognition load, we adopted a strategy similar to that used by Podd (1990) and Metzger (2002). The stimulus sets were constructed so that the 20 faces shown during the exposure phase for the LML group were identical to the first 20 of the 40 faces shown during the exposure phase for the HML group and identically ordered. Similarly, the 40 faces shown during the recognition phase for the LML group were the first 40 of the 80 faces shown during the recognition phase for the HML group. Therefore, the first half of both phases for the HML group was identical to the entirety of both phases for the LML group. As such, up until the 40th trial of the recognition task (the final trial for the LML group), the only difference between the two groups was that the HML group had experienced a greater memory load in the exposure phase. The rationale for this was that a comparison of recognition scores for the first 40 recognition trials would enable the effects of memory load to be assessed independently of the potentially confounding effects of recognition load (Podd, 1990). If memory load is the cause of the effect, the groups should exhibit differential accuracy; if, on the other hand, recognition load is the true cause, there should be no difference in accuracy between the groups.

The stimuli were 80 full-frontal color photographs of male faces, presented using a laptop computer with a 30-cm color screen. For all photographs, the model stood in front of a white screen and assumed a neutral expression. No potential memory cues other than the faces themselves were present; jewelry and glasses were removed, clothing was obscured, and none of the models had prominent features such as facial hair. Half the photographs were of male college students in their early to mid-twenties, and half were of older males between 63 and 97 years of age ($M = 78$, $SD = 7.48$). Within each target face age group (young and old), half the photographs were randomly assigned to be targets and half to be distractors. Therefore, there were 40 target photographs (20 young faces and 20 old) and 40 distractors (20 young and 20 old). Photographs of male faces were used to eliminate confounding by a sex of participant $\times$ sex of face interaction.

Procedure

The participants were tested in their own homes. A visual acuity test was given to ensure that they were able to see the faces on the computer screen clearly. Each participant was asked to read sentences in 18-point font from *Reading Test Types* as approved by the Faculty of Ophthalmologists, London (1987). The participants were seated comfortably at a table on which the laptop computer was placed. The screen was approximately 0.5 m away from the participant's face. Photographs of the target faces were presented one at a time during both the exposure and recognition phases. Each target face measured 12 cm $\times$ 15 cm and was centered on the computer screen. Each stimulus was presented for 5 sec, with an interstimulus interval (ISI) of 3 sec during which the screen was blank. There was no delay period between the exposure and recognition phases.

During the recognition phase, the participants indicated verbally whether they had seen the stimulus before. The participants responded "old" to photographs previously seen and "new" to photographs not previously seen. A forced choice procedure was utilized: the participants were instructed that, if they were unsure of the correct response, they should guess. During the recognition phase, a tone sounded at the end of each 5-sec stimulus presentation as a reminder to the participants that a decision had to be made, if it had not been already, within the 3-sec ISI. From the time each face appeared on the screen, the participants had 8 sec in which to respond. The LML group was exposed to 20 faces in the exposure phase and 40 in the recognition phase, whereas the HML group was exposed to 40 faces in the exposure phase and 80 in the recognition phase. When the recognition task was complete, the participants were thanked, debriefed, and informed of their results if they wished to know them.

RESULTS

The dependent variables for the present study were recognition accuracy (represented by d'), hit rate, false alarm rate, and c (a measure of response bias; Macmillan & Creelman, 1991). The d' statistic is particularly useful in face recognition research, for a number of reasons. Hit rate alone is unsuitable as a measure of recognition; after all, a perfect hit rate may be obtained by the artifice of saying "old" on all trials. It is therefore necessary to consider the false alarm rate as well. The d' statistic takes both hit rate and false alarm rate into account, and provides an estimate of recognition accuracy largely independent of any bias toward responding "new" or "old" (Macmillan & Creelman, 1991). We also analyzed hit rate and false alarm rate separately, in order to ascertain where the changes in d' occurred. When analyzing these variables, it is important to determine whether changes in either variable are due to genuine recognizability changes or to shifts in decision criterion. The c statistic was used for this purpose. Scores on the four dependent variables were analyzed with a series of ANOVAs within a $3 \times 2 \times 2$ mixed factorial design, with the two between-subjects variables of participant age (<40, 60–75, >75) and memory load (low, high) and the within-subjects variable of target age (young, old).

Hypothesis 1 stated that face recognition ability would decline with participant age. The data relevant to this hy-

Table 1
Means and Standard Deviations for Recognition Accuracy (d'), Hit Rate (HR), and False Alarm Rate (FAR) as a Function of Participant Age and Target Age

		d'		HR		FAR	
Participant Age	Target Age	*M*	*SD*	*M*	*SD*	*M*	*SD*
<40	All	2.14	0.80	0.79	0.11	0.15	0.11
	Young	2.03	0.88	0.79	0.12	0.17	0.12
	Old	2.24	0.71	0.79	0.11	0.12	0.08
60–75	All	1.90	0.77	0.80	0.15	0.22	0.13
	Young	1.59	0.69	0.76	0.15	0.26	0.14
	Old	2.21	0.73	0.84	0.13	0.18	0.11
>75	All	1.70	1.00	0.75	0.18	0.25	0.17
	Young	1.11	0.67	0.69	0.19	0.33	0.17
	Old	2.29	0.94	0.80	0.17	0.16	0.12

pothesis are presented in Table 1. There was a main effect of participant age on d' [$F(2,90) = 4.08$, $p = .02$, $\eta^2 = .08$, SP = .71].[1] Consistent with our expectations, the young participants performed at a higher level than did the oldest participants, with participants of ages 60 to 75 falling in between. The mean difference in accuracy between the youngest and middle groups is very similar to that between the middle and oldest groups. At first glance, this might appear to suggest a uniform rate of decline in recognition accuracy. However, the mean *age* difference between the youngest and middle groups (40.91 years) is considerably larger than that between the middle and oldest groups (14.38 years). Therefore, the results suggest an accelerating decline in the older participants.

The index d' was broken into its component parts—hit rate and false alarm rate—to determine where the changes in d' occurred. The effects of participant age on hit rate did not reach significance [$F(2,90) = 1.91$, $p = .15$, $\eta^2 = .04$, SP = .39]. However, for false alarm rate, a main effect of participant age was found [$F(2,90) = 8.28$, $p = .001$, $\eta^2 = .16$, SP = .96]. The effect size (η^2) for false alarms was four times the value for hits. Thus, the decreased face recognition accuracy associated with age was a product of an increased false alarm rate rather than a decreased hit rate, as predicted.

Hypothesis 2 stated that older people would be more accurate in recognizing old faces than young faces, whereas younger people would be more accurate in recognizing young faces as opposed to old faces. The main result relevant to this hypothesis is the significant interaction between participant age and target age for d' [$F(2,90) = 9.36$, $p < .001$, $\eta^2 = .17$, SP = .97]. This interaction qualifies the main effect of participant age on d' reported above, as well as the main effect of target age on d' [$F(1,90) = 53.98$, $p < .001$, $\eta^2 = .38$, SP $\approx$ 1.00]. Although the interaction is significant, the results do not conform wholly to our expectations. As Figure 1 shows, older adults exhibited poorer recognition of younger faces than of older faces (as predicted), but younger adults appeared to exhibit similar levels of recognition accuracy regardless of the age of the target face.

To corroborate this interpretation of the interaction, simple main effects were examined using *t* tests. For participants under 40 years of age, there was no statistically

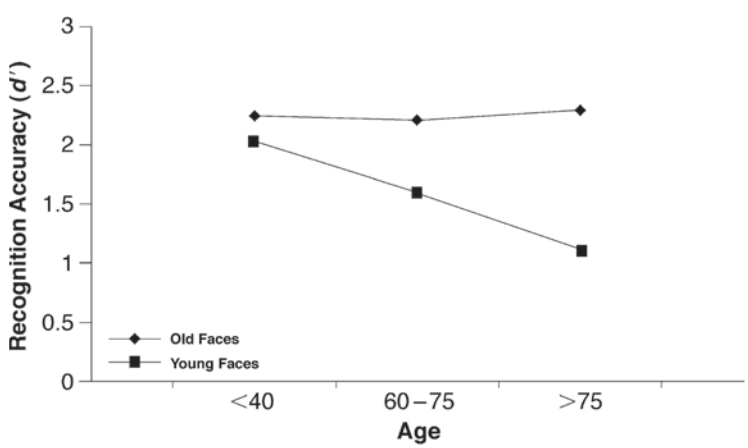

Figure 1. Interaction between participant age and target age for recognition accuracy (d').

significant difference in recognition accuracy for young and old target faces [$t(62) = 1.08, p = .28, d = .27$, SP = .17].[2] In contrast, participants 60 to 75 years of age were significantly worse at recognizing young faces as opposed to old faces [$t(62) = 3.68, p < .001, d = .87$, SP = .91], and participants older than 75 showed an even greater decline in recognition accuracy for the young faces than for the old faces [$t(62) = 5.82, p < .001, d = 1.44$, SP ≈ 1.00]. Thus, older adults performed progressively worse with younger faces, whereas younger adults were equally adept at recognizing younger and older faces. The participant age × target age interaction qualifies the conclusion that older adults exhibit poorer face recognition accuracy than younger adults. This conclusion applies only to younger faces.

The same interaction pattern found for d' was also evident when separate analyses were performed for hit rate and false alarm rate (see Table 1). First, there was an interaction between participant age and target age for hit rate [$F(2,90) = 3.86, p = .03, \eta^2 = .08$, SP = .69], qualifying the main effect of target age [$F(1,90) = 14.29, p < .001, \eta^2 = .14$, SP = .96]. In addition, there was an interaction between participant age and target age for false alarm rate [$F(2,90) = 5.97, p = .004, \eta^2 = .12$, SP = .87], qualifying the main effect of participant age reported above and the main effect of target age [$F(1,90) = 44.98, p = .001, \eta^2 = .33$, SP ≈ 1.00]. Note that the effect size for the false alarm rate is greater than that for the hit rate, which indicates that the age-related decrease in recognition accuracy for young faces was due more to an increase in false alarm rate than to a decrease in hit rate.

Hypothesis 3 stated that memory load would impair recognition performance and that the impairment would be greater for elderly participants. Table 2 shows the data relevant to this hypothesis. There was a main effect of memory load on d' [$F(1,90) = 6.35, p = .01, \eta^2 = .07$, SP = .70]. As expected, the level of accuracy found in the HML condition was lower than that obtained for the LML condition. There was no statistically significant interaction between memory load and participant age for d' ($F < 1$). This suggests that increased memory load was associated with a performance decrement in all groups of participants, irrespective of age, and appears to indicate that increased memory load had a small but consistent effect on recognition accuracy, as we predicted. Memory load had no consistent effect on hit rate ($F < 1$) but did produce a main effect for false alarm rate [$F(1,90) = 3.91, p = .05, \eta^2 = .02$, SP = .50]. This indicates that, once again, the decline in recognition accuracy was more a product of an increased false alarm rate than of a decreased hit rate.

However, to determine whether these findings were genuinely a product of memory load and not confounded by recognition load, the LML and HML groups were compared on the first 40 trials of the recognition task. As discussed earlier, if the difference in recognition accuracy between the LML and HML groups were due only to memory load, a significant effect should emerge in this comparison. The mean level of accuracy (d') for the HML group ($M = 1.98, SD = .96$) was lower than that for the LML group ($M = 2.07, SD = .96$); however, this difference was not significant ($F < 1$). This indicates that the difference between the LML and HML groups reported above was largely a product of recognition load, rather than memory load.

Finally, the c statistic, a measure of response bias, was calculated for each participant and aggregated across groups. The goal was to determine whether the differences in recognition accuracy found between ages were a product of differing levels of recognition accuracy or the product of response bias. If the latter were the case, values of c would deviate from 0, the value for a neutral criterion. As Table 3 shows, there was a relatively even spread of positive and negative criterion values (7 positive values vs. 5 negative values), with 7 of the 12 values deviating from 0 by .10 or less. Furthermore, for the larger mean deviations, the standard deviations were generally much larger than the mean deviation score itself, indicating that the mean deviations are unlikely to be of any importance. A series of ANOVAs confirmed that there were no statistically significant main effects or interactions for c. In short, the notable increase in false alarm rate and the smaller decrease in hits do not appear to be due to a change in response bias. Rather, aging seems to be accompanied by a genuine change in recognition accuracy.

DISCUSSION

Many of our hypotheses were confirmed, but there were also a number of surprises. As expected, greater age

Table 2
Means and Standard Deviations for Recognition Accuracy (d'), Hit Rate (HR), and False Alarm Rate (FAR) as a Function of Memory Load and Participant Age

		d'		HR		FAR	
Memory Load	Participant Age	*M*	*SD*	*M*	*SD*	*M*	*SD*
Low	All	2.07	0.96	0.78	0.17	0.19	0.14
	<40	2.38	0.86	0.81	0.12	0.13	0.09
	60–75	2.03	0.85	0.81	0.17	0.23	0.14
	>75	1.81	1.09	0.73	0.20	0.21	0.15
High	All	1.75	0.76	0.78	0.13	0.22	0.15
	<40	1.90	0.66	0.77	0.10	0.17	0.12
	60–75	1.77	0.67	0.79	0.12	0.22	0.13
	>75	1.59	0.91	0.77	0.17	0.28	0.18

Table 3
Means and Standard Deviations for *c*, as a Function of Participant Age, Target Age, and Memory Load

		Participant Age					
		<40		60–75		>75	
Target Age	Memory Load	*M*	*SD*	*M*	*SD*	*M*	*SD*
Young	Low	.09	.40	−.07	.59	.13	.47
	High	.10	.37	−.01	.40	−.23	.50
Old	Low	.16	.51	−.20	.57	.02	.64
	High	.21	.33	−.01	.33	−.04	.52

was associated with decreased face recognition accuracy. However, despite an overall decline in accuracy among older participants, this decline was not equally distributed across stimuli: It occurred for young faces but not for old faces. This result provides an important qualification to the claim that recognition accuracy declines with age. In addition, the result is consistent with our prediction that older adults would perform better with old faces than with young faces. However, the other half of our prediction—that young adults would perform better with young faces than with old faces—was not borne out. Young adults were equally adept at recognizing young and old target faces.

Our result was the reverse of what was found by Fulton and Bartlett (1991), who reported that recognition accuracy in their young adult groups was generally higher with young faces than with older faces, whereas recognition accuracy in the elderly groups showed no significant age effects. It is also inconsistent with Bäckman (1991), who found evidence that young adults are better at identifying young faces than old faces. Given that the present results conflict with previous research, no definite conclusion is possible. It is worth noting, however, that our result is consistent with those produced in laboratory crime simulations. For instance, Yarmey (1984) reported not only that elderly participants exhibited lower recognition accuracy than did younger participants, but also that recognition accuracy for elderly participants was lower with young faces than with old faces. It is possible that elderly participants have less processing resources available and apply these resources conservatively to the stimuli that are most salient to them: individuals of a similar age. In contrast, young participants have more processing resources and are therefore capable of retaining memories of faces even when those faces belong to members of groups that are less salient to them.

A comparison of hits and false alarms across the three age groups indicated that, although there was a small decline in hit rate for young stimulus faces, the overall decrease in accuracy was due largely to an increase in false alarms. This result is consistent with past research (Bartlett & Leslie, 1986; Crook & Larrabee, 1992; Ferris et al., 1980; Smith & Winograd, 1978; Yarmey, 1984). It is not immediately apparent how this finding can be explained. However, a clue may be found in the fact that it is not only aging that produces this pattern. Podd (1990) showed that increasing the length of the interval between the exposure and recognition phases in a face recognition task led to a decline in recognition accuracy, and that this decline was mainly the result of an increased false alarm rate. Furthermore, Davies, Shepherd, and Ellis (1979) showed that increasing the degree of similarity between target and distractor faces decreased recognition accuracy, again through an increase in the false alarm rate. As with the decline in face recognition accuracy with aging, the hit rate remained relatively constant in both cases.

What do these three circumstances have in common that could account for the increase in false alarms but unchanging hit rate? One possibility is that, in each case, fewer distinctive facial features are available to participants to make the judgment of whether or not they have seen the stimulus face before. With increasing age, details of the faces fade faster from memory (or perhaps are not encoded in the first place). With increasing retention intervals, there is more time for people's memories of the target faces to fade, with the least salient features fading fastest (Podd, 1990). The situation with increasing target–distractor similarity is slightly different. The problem here is not that there are fewer facial features stored in memory, but that the features that are in memory are so similar to those of the distractor faces that they are not very helpful in judging whether the face has been seen before. What all three cases have in common, therefore, is that participants have fewer *distinctive* facial features available in memory to make the judgment. Either the distinctive features are not in memory, or the features in memory are not distinctive.

The next question is, Why might the availability of fewer distinctive facial features lead to an increase in false alarms but not in hits? To simplify the explanation, consider two hypothetical participants. Participant A has only one distinctive facial feature encoded in memory; Participant B has five. With each face presented in the recognition phase, there are two possibilities: The face is a target, or it is a distractor. If it is a target face, both individuals have a reasonable chance of recognizing it—after all, accurate recognition of a target can occur with as few as one salient feature being available for the judgment. Therefore, in recognizing target faces, Participant A is not particularly disadvantaged in relation to Participant B, and their hit rates should be relatively similar. The situation changes when the stimulus face is a distractor. If one of the five facial features in Participant B's memory happens to be very similar to that of the stimulus face, Participant B can still compare the other features and may discern that it is a distractor. In contrast, if the one distinctive facial feature in Participant A's memory is very similar to that of the distractor, Participant A cannot check it against other features and is therefore more likely than Participant B to judge that it is a target face. In other words, Participant A's false alarm rate will be greater than that of Participant B. This theory may explain why a higher false alarm rate but a relatively stable hit rate occurs with increasing age, as well as why it occurs with increasing retention intervals and increasing target–distractor similarity.

Another important but unexpected finding of the present study relates to memory load. Although we found the usual association between increased memory load and decreased recognition of faces, our data suggest that the decrease in recognition accuracy was not in fact a product of memory load but of recognition load. Normally, these two factors are confounded, but when we separated them out, we found recognition load to be the true source of the effect. Few studies dealing with memory load have taken account of this potential confound, and our result challenges the interpretation of all such research. Effects that appear to be due to memory load may really be due to recognition load. A further finding of the present study was that the performance decrement due to recognition load was roughly uniform across age groups, indicating that recognition load affects older and younger participants in much the same way. This is consistent with the findings of Metzger (2002) and indicates that his result generalizes to older participants. Both recognition load and memory load are important factors to take into account in making comparisons across studies of face recognition accuracy.

The results of the present study may have a number of implications. One real-world implication concerns eyewitness testimony. Given the increased false alarm rate for older adults, elderly witnesses might be more likely to identify an innocent person as the criminal than might young adult witnesses. This potential problem is compounded by the fact that the false alarm rate of elderly adults is greater for young target faces than for old target faces (Yarmey, 1984). The present study also has implications for the conduct of future research in face recognition. Some authors (e.g., Crook & Larrabee, 1992) do not report the ages of their target faces. Our results suggest that this information is crucial for the proper interpretation of face recognition research. Furthermore, our results suggest that future theories of face recognition need to consider the interaction between the age of the participant and that of the target face. Statements about face recognition that might be true for one group of participants and one set of target faces may not be true for another. For instance, past studies have reported a generalized decline in face recognition accuracy with age, but the present study suggests that this decline may not apply to the recognition of old faces, which highlights the importance of focusing on possible interactions between participant variables and target face variables.

Finally, a few qualifications and cautions are necessary. First, it would have been desirable to measure and statistically control for potentially confounding variables such as verbal intelligence and vocabulary. This would have enabled us to rule out alternative explanations for the apparent age effects uncovered in this study (see Salthouse, 2004). Second, in common with many other studies, only male target faces were used in the present study. Although this eliminates the potentially confounding effects of target sex, it also limits the external validity of the study. Further research would ideally expand the present work through the inclusion of female target faces of differing ages. In addition, our sample contained three times as many females as males, which again could limit the external validity of the study. A more even balance of males and females would provide an improved test of the influence of aging on face recognition.

Conclusion

The present study utilized signal detection methodology to examine the effects of age of participant, age of target face, and memory load on face recognition accuracy. The study provides further empirical support for the view that decrements in face recognition accuracy increase with advancing age and that these decrements are more a product of increased false alarms than of decreased hits. In addition, the study makes several novel contributions. One important result is the finding that the decline of face recognition accuracy with age applies to young faces but not to old faces. Young adults seem able to distinguish between target and distractor faces equally well regardless of the age of the target, whereas older adults show progressively worse recognition of younger faces than older ones. The results of the present study confirm those of Podd (1990) in showing that increased memory load is associated with a reliable decrement in performance in recognition accuracy. However, our data suggest that memory load is confounded with recognition load and that it is recognition load that produces the decrement, which is independent of age. That is, the effects of increased recognition load seem to be constant across the life span.

REFERENCES

Bäckman, L. (1991). Recognition memory across the adult life span: The role of prior knowledge. *Memory & Cognition*, **19**, 63-71.

Bartlett, J. C., & Fulton, A. (1991). Familiarity and recognition of faces in old age. *Memory & Cognition*, **19**, 229-238.

Bartlett, J. C., & Leslie, J. E. (1986). Aging and memory for faces versus single views of faces. *Memory & Cognition*, **14**, 371-381.

Brigham, J. C., & Barkowitz, P. (1978). Do "they all look alike"? The effect of race, sex, experience and attitude on the ability to recognize faces. *Journal of Applied Social Psychology*, **8**, 306-318.

Carey, S. (1992). Becoming a face expert. In V. Bruce, A. Cowey, A. W. Ellis, & D. I. Perrett (Eds.), *Processing the facial image: Proceedings of a Royal Society discussion meeting held on 9 and 10 July, 1991* (pp. 95-103). Oxford: Oxford University Press, Clarendon Press.

Crook, T. H., & Larrabee, G. J. (1992). Changes in face recognition memory across the adult life span. *Journal of Gerontology*, **47**, 138-141.

Damasio, A. R. (1989). Neural mechanisms. In A. W. Young & H. D. Ellis (Eds.), *Handbook of research on face processing* (pp. 405-425). Amsterdam: North-Holland.

Davies, G. M., Shepherd, J. W., & Ellis, H. D. (1979). Similarity effects in facial recognition. *American Journal of Psychology*, **92**, 507-523.

Faculty of Ophthalmologists, London (1987). *Reading test types*. London: Society of Ophthalmologists.

Feinman, S., & Entwisle, D. R. (1976). Children's ability to recognize other children's faces. *Child Development*, **47**, 506-510.

Ferris, S. H., Crook, T., Clark, E., McCarthy, M., & Rae, D. (1980). Face recognition memory deficits in normal aging and senile dementia. *Journal of Gerontology*, **35**, 707-714.

Fulton, A., & Bartlett, J. C. (1991). Young and old faces in young and old heads: The factor of age in face recognition. *Psychology & Aging*, **6**, 623-630.
Goldstein, A. G. (1977). The fallibility of the eyewitness: Psychological evidence. In B. D. Sales (Ed.), *Psychology in the legal process* (pp. 223-247). New York: Spectrum.
Goldstein, A. G., & Chance, J. E. (1981). Laboratory studies of face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 81-104). London: Academic Press.
Lavrakas, P. J., Buri, J. R., & Mayzner, M. S. (1976). A perspective on the recognition of other-race faces. *Perception & Psychophysics*, **20**, 475-481.
Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
Metzger, M. M. (2002). Stimulus load and age effects in face recognition: A comparison of children and adults. *North American Journal of Psychology*, **4**, 51-62.
O'Toole, A. J., Peterson, J., & Deffenbacher, K. A. (1996). An "other-race effect" for categorizing faces by sex. *Perception*, **25**, 669-676.
Parkin, A. J. (1993). *Memory: Phenomena, experiment, and theory*. Oxford: Blackwell.
Podd, J. (1990). The effects of memory load and delay on face recognition. *Applied Cognitive Psychology*, **4**, 47-60.
Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, **13**, 140-144.
Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods*, **1**, 115-129.
Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, **100**, 139-156.
Shepherd, J. [W.] (1981). Social factors in face recognition. In G. Davies, H. Ellis, & J. [W.] Shepherd (Eds.), *Perceiving and remembering faces* (pp. 55-80). London: Academic Press.
Shepherd, J. W., Deregowski, J. B., & Ellis, H. D. (1974). A cross-cultural study of recognition memory for faces. *International Journal of Psychology*, **9**, 205-212.
Smith, A. D., & Winograd, E. (1978). Adult age differences in remembering faces. *Developmental Psychology*, **14**, 443-444.
Yarmey, A. D. (1984). Age as a factor in eyewitness memory. In G. L. Wells & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 142-154). Cambridge: Cambridge University Press.

NOTES

1. Post hoc statistical power (SP) levels were calculated using an alpha level of .05.

2. Note, though, that the level of power is low. It is possible that, with a greater number of participants, we would have detected a small but significant difference (see Schmidt, 1996).

(Manuscript received May 14, 2004;
revision accepted for publication September 18, 2004.)