Fortysomething:
Recognizing faces at one's 25th reunion

MAGGIE BRUCK
McGill University, Montreal, Quebec, Canada

PATRICK CAVANAGH
Harvard University, Cambridge, Massachusetts

and

STEPHEN J. CECI
Cornell University, Ithaca, New York

Twenty-four to 26 years after their high school graduation, subjects matched 40 photographs of former classmates taken from the high school yearbook with 40 current photographs of these same former classmates. Compared with age-matched control subjects who had no prior familiarity with the persons depicted in the photographs, the former classmates were superior at matching the old and new photographs. Generally, same-orientation photographs (i.e., old and new photographs taken from the same side) were matched more accurately than different-orientation photographs. Prior familiarity, however, rendered the orientation factor unimportant. These findings are discussed in terms of the processes that underlie facial recognition over very long intervals and in terms of their psycholegal implications.

Although many studies have examined facial recognition over short intervals (hours or even days) between initial presentation and test presentation of the faces, little is known about the durability of facial memory following very long intervals. Harry Bahrick and his associates reported the first of a series of studies of memory for high school classmates' yearbook pictures (Bahrick, Bahrick, & Wittlinger, 1975). They found that although adults had experienced some difficulty generating names for high school classmates, they were capable of recognizing names and pictures of classmates—as well as matching names with pictures of classmates—even up to 50 years after graduation.

Despite the immense theoretical importance of this work, it has had little influence in the practical and legal spheres of eyewitness facial memory. This is because, in Bahrick et al.'s task, recognition of former classmates' high school photographs requires only that a name be retrieved to match the static trace of the face as it was encapsulated in the high school photograph. To account

for such recognition, one does not need to posit any new mnemonic mechanisms, because much is already known about the generation and recognition processes that support this type of matching (e.g., Bahrick, 1970; Bruce & Young, 1986). These findings do not bear on the more practical issue that occasionally confronts courts of law—that of how difficult it is to recognize a current picture of someone who was known many years ago and whose appearance has changed during that time.

The well-publicized trials of accused World War II criminals Klaus Barbie ("The Butcher of Lyon") and John Demjanjuk ("Ivan the Terrible") are the most recent and celebrated instances in which defendants were identified on the basis of witnesses' memories of the defendant's appearance many years before. In the Klaus Barbie case, a French television station broadcast a photograph taken in the 1970s of a man living in Bolivia under the name of Klaus Altman. Viewers were merely asked if they recognized the man in the photograph, with no cue provided as to his identity. Many viewers from the Lyon region identified the man as Barbie, prompting the French government to extradite Barbie for "crimes against humanity." In the "Ivan the Terrible" case, the State of Israel accused John Demjanjuk of torturing and murdering internees at the Treblinka concentration camp. Demjanjuk, a factory worker who immigrated to the United States in the 1950s, denied that he was Ivan or that he had ever worked in any death camps. The legal proof of Demjanjuk's identity as Ivan was based on survivors' identifications taken approximately 35 years after their contact with Ivan. They were shown an array of pictures of known

or suspected war criminals and were asked if one of the pictures was Ivan's; one of the pictures was Demjanjuk's 1951 immigration photograph. All 5 courtroom witnesses selected Demjanjuk's photograph, and all 5 identified Demjanjuk in the courtroom as Ivan. However, according to Wagenaar (1988), these recognitions were not clear evidence of Demjanjuk's identity, because, among other factors, many survivors who did not testify could not correctly identify Demjanjuk's 1951 photograph. Of the 16 survivors questioned by Israeli authorities, 8 were unable to select the correct photograph, including Schlomo Helman, the survivor who had worked closest and longest with the real Ivan. Also, it is simply not known how many survivors were questioned, some estimate that the number was closer to 29, with only 8 positive identifications. The court discounted these inconsistencies, along with the expert testimony of the defense, because it felt that it was impossible for a survivor to err when faced with his torturer (Wagenaar, personal communication, 1988). Demjanjuk was sentenced to be executed in 1988, and the case is on appeal. Although little is known about the accuracy of this type of long-term recognition, the survivors' positive identifications have been challenged by the expert testimony of memory researchers, who, without exception, have relied on a corpus of eyewitness-memory research that is of only indirect relevance to the type of facial recognition required at these trials (Wagenaar, 1988).

One can add to these notorious World War II cases those instances in which claimants to fortunes showed up after long absences and aroused mixed opinions regarding their legitimacy. Ellis (1988) describes one such instance in British case law concerning the 19th century "Tichborne Claimant." According to Ellis (1988), members of the Tichborne claimant's immediate family, who had not seen him since he left England 13 years earlier in 1853 (and was subsequently reported to have drowned off the coast of Cuba), met his return with mixed views about his authenticity. Upon seeing him for the first time in 13 years, the claimant's mother said spontaneously, "Oh my dear Roger, is it you?" Although Lady Tichborne became the fifth person (in addition to two former servants, a cousin, and the family solicitor) to positively identify the claimant as the rightful heir to her late husband's estate, others (including his former tutor, several relatives, and former friends) were equally confident that the claimant was not Roger Tichborne. After two trials, the claimant was deemed to be an imposter and was imprisoned.

The Claimant of Tichborne case, the Klaus Barbie case, the Ivan the Terrible case, and the countless less celebrated ones (e.g., reports to national hotlines for child abductions that are based on matches of photographs taken several years ago and printed on milk containers with a face that is seen contemporaneously) reveal the difficulties and uncertainties expressed by some eyewitnesses, and the confidence of others, when they make contemporary identifications of faces they have not seen for long periods of time. The primary purpose of the present research was

to provide some empirical evidence about individuals' ability to identify faces from their distant past. Although not couched in a forensically relevant context, such as the cases described above, the present study was designed to provide a first pass at a problem for which currently there are only sparse data available.

In the present study, our subjects saw two sets of photographs—one containing photos of people's faces taken 25 years earlier and the other containing recent photos of the same persons. Each subject's task was to match the person in the past photo with that same person in the current photo. Some of the subjects had been friendly with one or more of the persons in the photographs, but they had not seen them for many years. Other subjects had no prior familiarity with the persons depicted in the photographs. Comparisons of these subjects' performance on the task not only provide data on the accuracy of long-term recognition of former acquaintances' faces but also inform us of some of the processes that are used for such recognition.

If facial identification after long lapses of time depends solely on perceptual processes that involve detecting facial features that survive the passage of time (e.g., matching the eye regions or the nose regions in the two sets of photos, searching for equivalent overbites or dimples in the two sets) or on the application of aging transforms (e.g., imaging wrinkles and receding hairlines) to the original photos (Enlow, 1982; Mark, Todd, & Shaw, 1981), then strangers should be similar to one-time friends in their ability to match past with present photographs, because there is no need to go beyond the information provided in the old and new photos (see Bruce & Young, 1986, for a discussion of the information used for face processing). Moreover, if the task can be performed solely on the basis of perceptual processes, then both groups should be equally affected by the manipulation of perceptual cues (e.g., orientation of the face in the photographs). Thus, it should be easier to match pictures that share the same facial orientation than it is to match pictures that differ in facial orientations (e.g., Bruce, 1982; Patterson & Baddeley, 1977).

But if this task can call on information beyond that contained in the actual photographs, then prior familiarity should convey specific advantages. That is, one-time friends might be familiar with different views of a classmate's face, as well as with characteristic expressions that are not represented in the high school photograph but which may be present in the current one. They may be able to directly recognize their former friends' faces in the current photos without their having to actually compare them with the high school photos to make a match. If such processes are important contributors to facial recognition after long intervals, then one-time friends should perform better on this task than strangers. So, although both friends and strangers could rely on perceptual strategies, including constructive ones, only the friends have the advantage of extrapictorial information. If strangers rely on perceptual information more than do

one-time friends, then they should be assisted to a greater degree by certain types of perceptual cues, such as similarity of orientation between the old and new photographs.

To examine these possibilities, we tested graduates of Westmount High School (WHS), a public school in Montreal. In May 1988, graduates of the Classes of 1962, 1963, and 1964 attended their 25th (plus or minus one) high school reunion. After the reunion, classmates who did not attend were asked to serve as subjects in this study. They were asked to match pictures of their classmates, taken from their high school yearbooks, to current photographs of these same classmates. Their responses were compared with those of age-matched subjects who had never lived in the Montreal area.

METHOD

Stimuli

High school photographs. Pictures of 40 classmates (20 male, 20 female) were taken from the high school yearbooks of the Classes of 1962, 1963, and 1964. All pictures were in black and white and were head-and-shoulder views of each person. The orientations of the faces were either three-quarter right or three-quarter left. The average age of the classmates was 16 years at the time the photographs were taken (1962–64). The size of each graduating class was approximately 210 students.

Follow-up photographs. Approximately one month before the 25th reunion, 60 classmates (30 males and 30 females) who lived in the Montreal area were photographed. These 60 classmates included the 40 whose yearbook pictures had been selected, as well as 20 others, who served as classmate foils. The classmates were in their early forties at the time of the experiment. In addition, 20 photographs were taken of 40-year-olds (10 male and 10 female) who did not attend WHS and who were professional colleagues of two of the authors. They were selected because they were not likely to have met or to have been familiar to any of the WHS graduates.

The subjects were posed to resemble, as much as possible, the high school yearbook photographs. The pictures were head-and-shoulder views with orientations either three-quarter right or three-quarter left. However, for the present set of pictures, as opposed to the high school pictures, the subjects removed their eyeglasses. Photographs were taken with a Polaroid camera and flash. Subjects sat in front of a gray backdrop (similar to the one used in the graduation pictures), 132 cm from the lens of the camera. They were told to turn their heads to the left or to the right, to look straight ahead, and to have neutral facial expressions. These photographs were then reshot in black and white and resized so that they would be comparable to the yearbook pictures.

Test booklets. Each booklet contained eight pages. On each page there were 5 high school pictures and 10 follow-up pictures (2.2 × 3.5 cm). The 5 high school pictures were placed vertically in the first column on the left-hand side of the page; the 10 follow-up pictures were placed in two columns of 5 on the right-hand side of the page. Each follow-up picture was numbered 1 through 10. Five of the follow-up pictures matched the 5 high school pictures in the left-hand column. The other 5 were photographs of other classmates and nonclassmates. On each page, there were either all males or all females. (See the Appendix for a sample page from the booklet.) Each picture appeared only once in the booklet. Thus, for each sex, there were four pages of photographs containing 20 high school pictures and 40 potential matches: 20 follow-up photos, 10 classmate foils, and 10 nonclassmate foils.

In making up the test booklets, we tried to arrange each page so that the five high school pictures were similar (e.g., all blonds) and so that the follow-up pictures and their foils were reasonably homogeneous. For 22 of the 40 matched pairs, the orientation of the faces (left- or right-facing) was the same in the high school and in the follow-up pictures, whereas for the remaining 18 faces, it was reversed.

Subjects were asked to name the person in each high school photograph and to indicate the number of the current photograph that best matched the high school picture. A three-point scale was used to ask how well that person was known by the subject during high school (very well, somewhat, not at all). Next, subjects were asked when (the number of years) they had last seen the person in the high school photograph. If the person was unknown, there was a space to check "never." Finally, under each of the current photographs, subjects were asked to place a mark in the box "recognize" if they thought they recognized any of the people on the page. (It should be noted that this last instruction was rarely followed and, for this reason, these data are not reported in this paper.)

The eight pages were arranged so that the sex of the classmates alternated on consecutive pages. In addition, there were two different page orderings of the booklet. A cover sheet that contained instructions and two demonstration photographs was stapled on the front. These demonstration photographs were not repeated in the test booklets.

Procedure

Several weeks after the high school reunion, the test booklets, along with a letter describing the purpose of the study and the instructions for completing the booklets, were mailed to 110 classmates who did not attend the reunion. The following instructions were included in the letter:

1. Match each of the 5 pictures in the left-hand column of the following pages with 1 of the 10 in the two right-hand columns of that page. Write the number of your choice where it says "Match to."
2. Try to guess the classmate's name.
3. Check off how well you knew this person in high school.
4. Write down the number of years it has been since you last saw this person.
5. Under *each* picture in the right-hand column, make a check in the box if you recognize this person. Note, it doesn't necessarily have to be a correct match with one of the five yearbook pictures.
6. Please complete the information for each picture on the page—even if you think you have never seen this person before, try to guess what he or she looks like now.
7. Don't take out your yearbook until you have finished doing all the pictures. Then if you decide to check your guesses, don't make any changes in the booklet.
8. Return the booklet in the self-addressed envelope.

Subjects

Of the 104 questionnaires sent to WHS graduates, 50 were returned. Ten were excluded from further analysis either because subjects did not correctly complete the test (the most general case was to complete items for familiar classmates only) or because subjects had only attended the high school for one year and were not familiar with any of the pictures. The average age of the 40 subjects included in this study was 42.1 years; exactly half of them were female.

For the control data, 40 subjects were recruited from New York State and had no prior familiarity with any of the persons depicted in the photographs, had not attended WHS, and had never lived in Montreal. These subjects were recruited from a combination of sources, including recreational centers and the university subject pool. They were selected to resemble the WHS graduates in age, race (all Caucasian), and gender. Their mean age was 42.3 years; half of them were female. They completed the matching task alone in a quiet testing room. The testing protocol was similar, though not identical, to that described above for the WHS graduates. Control subjects were instructed to take as long as they wished to make their matches; they took, on average, 31 min to finish the 40-photo booklet. They were asked to indicate if any of the persons depicted

in the photographs were familiar to them and, if so, from what context. None reported having seen any of the persons before. Finally, unlike the WHS graduates, the control subjects were asked to state their confidence for each match on a three-point scale (very confident, moderately confident, not at all confident) to provide a measure of comparability with other studies on facial recognition that have reported confidence–accuracy correlations when subjects were strangers.

RESULTS

For the WHS data, responses were excluded if the subject had seen a former classmate within the last 17 years (4.6% of all observations). This cutoff was selected because preliminary analyses revealed no reliable differences for correct responses if former classmates had not been seen over the 18–25-year period ($p > .10$). However, there were substantial differences between rates of correct responses when subjects had seen friends within the past 17 years compared with when they had not seen them for more than 18 years (87% vs. 65%, $\chi^2 = 14.26, p < .01$).

Data from both the WHS graduates and the control-group subjects were first analyzed in the same regression model. The independent variables were group (WHS graduates vs. New York State controls), sex of subject, sex of photograph, and orientation of old and new pictures (same vs. different). Both "individuals" and "photographs," nested within orientation, group, and sex of subject, were regarded as levels of random factors.[1] This resulted in a mixed model in which covariates could be assessed at both within- and between-level stages (e.g., the covariate "sex of photograph" could be examined at the level of the individual photograph and/or subject, as well as at the level of an entire orientation, sex of photograph, or sex of subject). Treatment of individuals and stimuli as random factors allows inferences to be made to a larger population of individuals and photographs than is used here and provides a powerful means of nesting and aggregating interaction effects.

The main effect for group was highly significant [$F(1,76) = 24.63, p < .0001$], despite pronounced individual variation within each group [$F(1,76) = 1.70, p < .05$]. The mean accuracy score for the WHS graduates was 49% (with ranges from 33% to 78%), whereas the control subjects' mean accuracy score was 33% (with ranges from 11% to 72%). For both groups (and all subjects), the rate of correct matching was significantly greater than chance [both $ps < .001$], for which the probability of matching old and new pictures is 10%. (The probability calculation is based on the finding that subjects responded without replacement; that is, a single current picture was never matched with more than one high school photograph.)

There was a reliable interaction involving sex of subject × sex of photograph [$F(1,76) = 4.07, p < .05$] that was qualified by a three-way interaction involving group × sex of subject × sex of photograph [$F(1,76) = 7.68, p < .01$]. Planned contrasts showed that the three-way interaction occurred because the sex of subject × sex of

photograph interaction was significant for the WHS graduates [$F(1,3113) = 16.43, p < .001$] but not for the control subjects ($F < 1$). The WHS subjects' rates for correctly guessing photographs of their same sex and opposite sex were 55% and 45%, respectively. The rates for the control subjects were 33% and 34%.

There was a marginal main effect of orientation [$F(1,76) = 2.66, p < .10$], and the interaction between group and orientation approached conventional levels of statistical significance [$F(1,76) = 3.27, p < .076$]. Planned comparisons showed that, as predicted, the orientation effect was significant for control subjects [39% for guessing photographs with similar orientations vs. 29% for guessing photographs with dissimilar orientations; $F(1,3113) = 6.70, p < .01$] but not for WHS subjects [51% and 49%, respectively; $F(1,3113) = 1.06, p = .23$].

Next, to explore the effects of confidence (for the control-group data) and familiarity (for the WHS data) on the above findings, a set of analyses was run separately on the data from each group. These analyses eliminated the classification effect of group but included all other factors that were entered into the mixed model analysis, as well as familiarity ratings for the WHS group and confidence ratings for the control group. Because of asymmetries in the proportions of photographs that were rated as highly familiar, moderately familiar, and unfamiliar, as well as asymmetries in the proportions of photographs that were rated as very confident, moderately confident, and not at all confident, it was necessary to run these analyses in a fixed model.

For the analysis of the control-group scores, there was a significant main effect for confidence [$F(2,1576) = 63.47, p < .001$]. The mean accuracy on the photos that were given the highest confidence ratings was 63%, whereas for the moderate- and low-confidence ratings the recognition accuracy means were 40% and 24%, respectively. Although these differences are of a large magnitude and are reliably different from each other (Fs ranging between 31 and 107 for the contrasts, all $ps < .001$, by single degree-of-freedom tests), there were such large variations within each confidence level that there was only a modest zero-order correlation between overall recognition accuracy and confidence ($r = .26, p < .01$). There was a significant orientation × sex of photograph interaction, which was qualified by orientation × sex of photograph × confidence [$F(2,1576) = 4.51, p = .012$]. Analysis of the three-way interaction showed that four of the six contrasts were in the predicted direction. Photographs with similar orientations were easier to match than were those with dissimilar orientations (the two contrasts that ran counter to this trend were female photographs at high levels of confidence and male photographs at moderate levels of confidence). Because we had no a priori hypotheses concerning this specific interaction, and because it is not reported elsewhere in the published literature (nor does it evoke intriguing post hoc theoretical explanations), we shall not discuss it further, as doing so might obfuscate the overall role of orientation in the control-group

data. Sex of subject did not emerge as a reliable main effect or as an interactive contributor to any of the control-group findings.

For the analysis of the WHS data, there was a significant main effect of familiarity [$F(2,1505) = 34.32, p < .001$]. Accuracy means for the photographs that were rated as very familiar, somewhat familiar, and unfamiliar were 78%, 58%, and 37%, respectively. (These values do not sum to the grand mean of 49% because of asymmetries in the proportions in each category.) Single degree-of-freedom significance tests showed that all of the above means differed significantly from each other (Fs ranging between 9 and 42, all ps $< .01$). Although these means are quite disparate, the large degree of within-category variability dampened the overall correlation between familiarity and accuracy ($r = .25, p < .01$). Interestingly, the accuracy rate for unfamiliar photographs among these WHS subjects is nearly the same as the control group's overall mean of 33%, suggesting that the WHS subjects were, in fact, truly unfamiliar with these pictures. It should also be noted that the subjects' ability to match unfamiliar photos was not significantly correlated with their ability to match familiar photos ($r = .32, p > .05$, 1-tailed test).

Familiarity interacted with orientation [$F(2,1505) = 3.33, p < .05$]. When the subjects reported that they had known the persons depicted in the photos either very well or moderately well, their accuracy rates were statistically equivalent for the similar and dissimilar orientations. On the other hand, when the subjects reported not having known the person in the photos, their accuracy rate was 8% higher for similar orientations [$F(1,1505) = 4.2$, $p < .05$].

Finally, when familiarity was controlled, there was no interaction between sex of subject × sex of photo, rendering the WHS results comparable to those of the control group in this regard. Thus, WHS subjects were more likely to recognize own-sex, as opposed to opposite-sex, photographs because they were more likely to have been friends with these own-sex classmates, not because of an inherent same-sex matching advantage.

It might be argued that the familiarity effects may have been due to a spurious correlation between perceptual similarity of old and new photographs and familiarity. However, two results rule against this hypothesis and suggest that the familiarity effects obtained for the WHS subjects were not confounded by the case of specific items. First, there was a nonsignificant correlation ($r = .06$), between each photograph's familiarity rating by the WHS subjects and how accurately it was recognized by the control-group subjects. Thus, the photographs that were rated as most familiar by the WHS subjects were not systematically the easiest items for the control group. Second, because individual photographs were treated as a random variable in the mixed model, the results reported earlier are adjusted for the individual characteristics of each photograph, including its familiarity rating.

In the last analysis, the "naming" data from the WHS items were examined. WHS subjects provided names for 21% of the high school photographs. When a name was provided, it was incorrect 11% of the time, partially correct (either first or last name was correct) 18% of the time, and totally correct (correct first and last name) 71% of the time. As one would expect, provision of correct or partially correct names was highly associated with familiarity ratings. The subjects provided correct or partially correct names for 87% of all pictures rated as very familiar, for 42% of all pictures rated as moderately familiar, and for only 2.5% of all pictures rated as least familiar. At each level of familiarity, the ability to correctly recall a part or all of a name increased the probability of correct matches. For very familiar items, the correct matching rates were 84% (when the correct name was given) versus 33% (when no name was given); for moderately familiar items, the comparable rates were 72% versus 50%; and for least familiar items, the rates were 64% versus 40%.

DISCUSSION

Subjects were much better at matching old to new photographs if they had once known the person in the photographs. This finding is consistent with some results of the face recognition literature—matches for two photos taken at the same historical point are more accurate for familiar than for unfamiliar faces (e.g., Young, McWeeny, Hay, & Ellis, 1986). In the case of the present study, familiarity involves, among other things, having formerly seen the persons depicted in the photographs in many guises (hair styles, skin tones), emotional expressions, and orientations other than those pictured in the high school photographs. This may allow the subjects who were, at one time, familiar with the persons depicted in the old photographs to recognize current photographs, not because these photographs look like the ones taken 25 years ago, but because they resemble or re-invoke some characteristic pose, expression, or guise of the subjects' former classmates that the subjects recall. It is also the case, however, that even those subjects with no former contacts with the people depicted in the photographs (i.e., control subjects and those WHS subjects who reported low levels of familiarity) performed significantly better than chance. This suggests that perceptual information alone can be used to perform this task with some level of proficiency.

Before we discuss other details of the results, it is important to first deal with the potential problem of subject selection bias for the WHS subjects, which, if present, could limit the generality of our results. The return rate for the test booklets was 48% for the WHS subjects. Although this falls well within the boundaries of most survey-research return rates, it raises the possibility that those who returned the booklets either knew the most classmates when they were in high school and/or were particularly good at this task. To whatever extent either of these

accounts is true, our data are "silent," other than showing (1) that our respondents were widely distributed across the "familiarity" spectrum, with many respondents reporting having known few of their former classmates, and (2) that there was much variation within the WHS group, with some subjects performing more poorly than control subjects. Moreover, there was also some selection bias among the control subjects themselves—not all control subjects who were asked to participate did so. We have no figures of overall participation rate among persons approached to be control subjects, but in one group of adults that was solicited, only 5, out of approximately 12, agreed to participate. Thus, one might argue that the control group's performance was higher than might be obtained if all who were solicited agreed to participate. Finally, the finding that the control group's overall accuracy rate was identical to that of the WHS subjects' accuracy rate for unfamiliar subjects suggests that the two groups of subjects were sampled from the same population. Thus, taken together, the 48% return rate of WHS subjects probably did not influence the major results reported.

Compared with WHS graduates, control subjects, as predicted, showed a significant orientation effect—they were more likely to correctly match photographs that shared the same facial orientation than they were pairs that had different orientations. Furthermore, WHS subjects who reported low levels of familiarity also did better when photos shared similar orientations. Both results support the hypothesis that perceptual cues play a greater role in the recognition of unfamiliar than of familiar faces and that extrapictorial information is used for the recognition of familiar photographs.

In addition to the benefit of having once known someone, for the WHS data there appeared to be greater accuracy when subjects matched photographs of their own sex than there was when they matched photographs of the opposite sex. At first, this effect is reminiscent of studies of cross-racial face recognition, which indicate that it is more difficult to remember faces of an unfamiliar race than it is faces of one's own race (e.g., Shepherd, Deregowski, & Ellis, 1974). In the present study, however, the sex of photo $\times$ sex of subject interaction was no longer significant when familiarity was entered as a covariate, because there was a strong tendency among former WHS classmates to be more familiar with members of their own gender than with members of the opposite gender, thus leading to the observed superiority at matching own-sex photos.

The naming data provide some support for the familiarity ratings. In most cases, when subjects said they knew a person very well, they supplied the name of that person. The ability to supply correct names decreased as a function of reductions in familiarity. Further analyses showed that, controlling for level of familiarity, subjects were more likely to make accurate matches if they could correctly supply the name of the person. These results are consistent with previous suggestions that names may be an independent code that facilitates the task of matching (Bruce & Young, 1986; Hanley & Cowell, 1988).

The confidence-accuracy relation found for our control subjects ($r = .26$) is very similar to the overall mean of .25 reported in a meta-analysis of 35 studies on face recognition (Bothwell, Deffenbacher, & Brigham, 1987), even though the subjects in the latter tasks were asked to recognize unfamiliar faces seen but once. Recently, Cutler and Penrod (1989) reported substantial variability around the mean confidence-accuracy relationship, with factors such as cue degradation, distinctiveness, and base rate information (e.g., how probable the match was) moderating the relationship. In explaining the variability in findings, these authors suggested that lower correlations between confidence and accuracy should be found when disguises or facial changes are used, a result we did not observe here, with naturally aged photos.

Caveat lector. In closing, we raise several caveats about the applicability of these findings to the legal arena. First, the present study conditions are quite unlike those that bring most victims/witnesses to testify in court in the "crimes against humanity" cases alluded to in the introduction to this paper. Those who testify in such cases usually have had extreme emotional experiences (Wagenaar, 1988). In contrast, it is unlikely that the WHS subjects' interactions with their former classmates who were photographed in this study were at the same level of emotional turmoil. It could be that high levels of stress that accompany some eyewitness conditions result in an actual deterioration of memory (Deffenbacher, 1983; Loftus & Burns, 1982; see, e.g., Wagenaar & Groeneweg's, 1990, report that many concentration camp survivors who had been tortured were unable to recognize the names or faces of their tormentors 40 years later). In addition, our subjects had more years of contact with the photographed classmates than courtroom witnesses generally have with their assailants. The factors of stress and time might explain why recognition was high for known classmates in our study (58% for moderately familiar and 78% for very familiar photographs) and why it was low (28% to 50% of questioned survivors made correct identifications) in the Demjanjuk case.[2]

Independent of the effects of stress and time factors on recognition memory, the present data probably underestimate individuals' ultimate ability to recognize persons from their past if these persons are presented in a "live" lineup rather than in a photographic context, as was done in the present study. This is because static photographs eliminate potentially valuable cues (linguistic, acoustic, kinesthetic, affective). Any of these cues could make the identification of a former acquaintance much easier than our task. Having said this, our data nevertheless would appear to be directly relevant to other photographic contexts, such as the recognition of missing children from a comparison of old and current photographs or the recognition of someone from a photograph on the FBI's "Ten Most Wanted" list in the local post office.

Finally, our data show the same pattern of intersubject inconsistencies as found in some of the psycholegal cases presented in the introduction (i.e., Ivan the Terrible and

the Tichborne claimant). Rarely do all witnesses or subjects concur on the accurate identification of someone from their past. The variation on our task reflects two factors which should be considered in evaluating courtroom testimonies and in future psycholegal research. The first concerns the degree of physical change between past and present appearances. There is enormous variability across the faces depicted in the photographs in how extensively they aged. Although there is no generally accepted means for coding and quantifying such changes, some faces are clearly more difficult to match than others, often because of extensive hair loss, weight gain, wrinkling, and so forth. Second, there is enormous variation across subjects: Some control subjects performed better than some WHS subjects, including ones who once knew the persons in the photographs quite well. Thus, not all witnesses may reliably recognize all defendants. Having noted such individual differences, however, we have little knowledge of the factors that contribute to them. The degree to which individual differences in recognition accuracy is a function of a person's career, of daily activities that involve craniofacial processing (e.g., illustrators, plastic surgeons), of various components of intelligence (e.g., visuospatial abilities), or of personality characteristics (e.g., field dependence) is unknown. The study of the role of such factors is of forensic, as well as theoretical, importance.

REFERENCES

BAHRICK, H. P. (1970). A two-phase model of prompted recall. *Psychological Review*, 77, 215-222.

BAHRICK, H. P., BAHRICK, P. O., & WITTLINGER, R. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General*, 104, 54-75.

BOTHWELL, R., DEFFENBACHER, K., & BRIGHAM, J. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revised. *Journal of Applied Psychology*, 72, 691-695.

BRUCE, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology*, 73, 105-116.

BRUCE, V., & YOUNG, A. (1986). Understanding face recognition. *British Journal of Psychology*, 737, 305-327.

CUTLER, B., & PENROD, S. (1989). Moderators of the confidence-accuracy correlation in face recognition: The role of information processing and base rates. *Applied Cognitive Psychology*, 3, 95-107.

DEFFENBACHER, K. (1983). The influence of arousal on reliability of testimony. In B. Clifford & S. Lloyd-Bostock (Eds.), *Evaluating witness evidence* (pp. 235-251). Chichester, England: Wiley.

ELLIS, H. D. (1988). The Tichborne Claimant: Person identification following very long intervals. *Applied Cognitive Psychology*, 2, 257-264.

ENLOW, D. H. (1982). *Handbook of Facial Growth*. Philadelphia: W. B. Saunders.

HANLEY, J. R., & COWELL, E. S. (1988). The effects of different types of retrieval cues on the recall of names of famous faces. *Memory & Cognition*, 16, 545-555.

LOFTUS, E. F., & BURNS, T. E. (1982). Mental shock can produce retrograde amnesia. *Memory & Cognition*, 10, 318-323.

MARK, L. S., TODD, J. T., & SHAW, R. (1981). Perception of growth: A geometric analysis of how different styles of change are distinguished. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 855-868.

PATTERSON, K. E., & BADDELEY, A. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning & Memory*, 3, 406-417.

SHEPHERD, J., DEREGOWSKI, J., & ELLIS, H. (1974). A cross-cultural study of recognition memory for faces. *International Journal of Psychology*, 9, 205-211.

WAGENAAR, W. A. (1988). *Identifying Ivan: A case study in legal psychology*. Cambridge, MA: Harvard University Press.

WAGENAAR, W. A., & GROENEWEG, J. (1990). The memory of concentration camp survivors. *Applied Cognitive Psychology*, 4, 77-87.

YOUNG, A. W., MCWEENY, K. H., HAY, D. C., & ELLIS, A. W. (1986). Matching familiar and unfamiliar faces on identity and expression. *Psychological Research*, 48, 63-68.

NOTES

1. In addition, two other factors were included in our initial analyses of these data—location on a page and page within the booklet. The former was included to examine differential accuracy rates as a function of a photograph's serial order on a page. The second factor was included to determine whether or not factors such as "fatigue" or "strategy shifts" occurred as subjects worked through the booklet of photographs. Due to their lack of significance, both as main effects and in interaction with other factors, these analyses are not reported and these factors were not included in subsequent analyses.

2. There are other factors that must be considered in directly comparing the recognition data of the classmates and of the survivors in the Demjanjuk case. The first is that the classmates were tested after 25 years, whereas the survivors were questioned after approximately 35 years. The factor of time of testing would increase the rates of correct response for the classmates. Second, the survivors were shown a photograph taken 10 years after their last contact with Demjanjuk, whereas the classmates were shown a photograph taken 25 years after their last contact with their friends. This factor should tend to make the task more difficult for the classmates than for the survivors. Therefore, on net, our findings raise potential concerns about the accuracy of the identifications in the Demjanjuk case.

APPENDIX