Zimbabwe Law Review

Volume 14 1997

THE ZIMBABWE LAW REVIEW

1997 VOLUME 14

CONTENTS

ARTICLES

CHILDREN AND ADULTS' MEMORY FOR FACES: IMPLICATIONS FOR FACE IDENTIFICATION PROCEDURES IN CRIMINAL INVESTIGATIONS

Patrick Chiroro, Stabiso Sithole and Tinashe Muromo

Lecturers, Department of Psychology, University of Zimbabwe

ABSTRACT

Witnesses to a crime are often required to identify a suspect from a police line-up or photo spread. Positive identification of a suspect made under oath and with confidence by one or more witnessess is usually considered to be strong evidence that the suspect is indeed the criminal. However, cases of mistaken identity sometimes do occur, often with disasterous consequences for the misidentified persons and their families (e.g. the cases of John Demjanjuk, Laszlo Virag, Frank Walus & Marinus de Rijke — see Wagenaar, 1988, pp 1–30). In recent years, applied research on the psychological processes involved in human memory has contributed significantly to our understanding of the sources of erroneous identifications in criminal investigations. However, research on developmental trends in face identification ability has been sparse. In this article, we report the results of two experiments in which children and adults's face recognition and identification performance were examined. Findings from both experiments showed that: (i) face recognition performance and face identification ability are skills that develop and improve as a function of increasing age, (ii) the probability of false recognitions occurring in a police line-up or photo spread is smaller for faces with distinctive features and/or feature configurations than it is for typical or 'average-looking' faces, and (iii) when the 'to-be-remembered' face is distinctive in appearance, children as young as 6–7 years are no more likely to make face identification errors than are adults. The practical implications of these findings for face identification procedures in criminal investigations are discussed.

INTRODUCTION

Positive identification of a suspect by one or more witnesses is often considered to be strong evidence that a suspect might indeed be the criminal. While in most cases, person identification evidence is corroborated by other forms of evidence leading to successful conviction, erroneous 'identifications' do sometimes occur, often with disastrous consequences for the mis-identified persons and their families. The much publicised trial of John Demjanjuk is a case in point. John Demjanjuk was tried in a special court in Jerusalem during the period 1987–1988. He was charged for being "Ivan the Terrible" who killed some 850 000 Jews in the gas chambers of Treblinka. The case against Demjanjuk rested purely on face identification evidence given by several holocaust survivors who served as witnesses almost 44 years after their initial encounter with "Ivan the Terrible".

The procedure followed in the identification process was as follows. Mrs Radiwker, a member of the Israeli Police started by publishing an advertisement in newspapers inviting survivors of Sobibor and Treblinka to contact her for questioning regarding atrocities committed by two Ukrainians: Ivan Demjanjuk and Fedor Fedorenko. Thus, the witnesses had prior information regarding the defendants' names before they went to see Mrs Radiwker. Upon arriving at the Israeli Police Headquarters, witnesses were first interviewed by Mrs Radiwker or Martin Kolar (later) before being shown photographs of John

Demjanjuk and Fedor Fedorenko together with some foils (distracters). The witnesses were asked to study the pictures and indicate any faces which they thought were of Nazi criminals. Although the prosecution argued that John Demjanjuk's picture was identified as being that of "Ivan the Terrible" by several witnesses, there were some serious flaws in the manner in which the identification evidence was gathered. Demjanjuk was eventually acquitted largely because of expert witness evidence given by a psychologist who has since written a whole book describing his ". . . reflection upon (his) testimony and the scientific knowledge on which it (the testimony) rested" (see Wagenaar, 1988 for a full account of Demjanjuk's case). Several similar cases of misidentifications have been reported in the literature (e.g. the case of Mr. Virag; the case of Frank Walus; and the case of Marinus de Rijke, *see* Wagenaar, 1988, pp. 15–30).

Police officers, prosecutors, lawyers, and judges need to acquaint themselves with the scientific evidence that psycho-legal researchers have accumulated regarding possible sources of misidentification in criminal cases. We strongly believe that such empirical findings must be incorporated into the evidence gathering and/or evaluation process instead of relying on conjecture, whim, or mere precedences. Recently, *The Herald* reported the case of one aspiring boxer in Zimbabwe who, as a result of misidentification, was arrested and heavily assaulted for having mugged a traveller at a bus terminus. His boxing career is now in ruins and might never be resuscitated. Psycho-legal research has shown that errors of identification can arise from several sources. These include perceptual factors that operate within the witness's cognitive system (*see* Wells & Turtle, 1987; Fruzzetti, Toland, Teller & Loftus, 1992; and Doris, 1995), situational/contextual factors that operate during the commission of a crime (*see* Clifford & Scott, 1978; Leippe, Wells & Ostrom, 1978; Brigham, Maas, Snyder, & Spaulding, 1982; and Fisher, Amador & Geiselman, 1989), and factors that are under the control of investigating officers such as the manner in which witnesses are interviewed and the procedures that are followed in constructing photo spreads and police line-ups (*see* Wells, Leippe & Ostrom, 1979; Malpass & Devine, 1981; Morris & Morris, 1985; Geiselman, Fisher, MacKinnon, & Holland, 1985; Wells, 1990; and Wright, 1993).

In this article, we describe and present results of two experiments that we recently conducted to examine the extent to which two variables affect recognition memory for faces and face identification evidence. In experiment 1, we investigated the effect of differences in age on recognition memory for 'distinctive' and 'typical' faces. 'Distinctive faces' were faces whose features and/or feature configurations were independently judged to be unique or unusual in appearance while 'typical' faces were faces with 'average-looking' features and/or feature configurations. In experiment 1, the standard recognition memory paradigm was used (see Murdock, 1982). In experiment 2, we tested different participants on their ability to identify face photographs of persons with whom they had had meaningful contact three months earlier. The primary goal in both experiments was to examine the main effects of age and of face distinctiveness as well as to establish the extent to which these two variables would interact in determining recognition accuracy.

EXPERIMENT 1

A 4 X 2 split-plot factorial design was used in this experiment. Age served as a between-subjects factor and face distinctiveness was a within-subjects factor. A total of 148 children and adults aged between 6 and 18 years participated in the experiment. Half of the participants (group 1) took part in a ranking exercise that was designed to enable us to classify the faces into equal numbers of 'distinctive' and 'typical' faces. The other half of the participants (group 2) took part in the 'face recognition memory experiment' part of the study. For each part of the study, participants fell into four age groups as follows: (i) 6–7 year olds, (ii) 9–10 year olds, (iii) 12–13 year olds, and (iv) 16–18 year olds. In this experiment, the latter group served as 'adults'.[1] All the participants were drawn from primary and secondary schools in one of Harare's low density suburbs over a period of 8 months. Face distinctiveness was manipulated by testing all group 2 participants on their recognition of equal numbers of 'distinctive' and 'typical' target faces. The dependent variables were: (i) proportion of hits, and (ii) proportion of false positives.

The concept of 'hits' and 'false positives' (also called false identifications) derives from a psycho physics tradition that is based on signal detection theory whose origins lie in statistical decision theory (see Green & Swets, 1966; and Macmillan & Creelman, 1991). In a recognition task, a 'hit' occurs when a participant (or witness) correctly identifies a presented stimulus (e.g. a face) as one seen previously while a false positive is, as the name implies, a positive response to a stimulus (face) not seen before. For the legal system, a hit is a welcome response that is likely to lead to a valid conviction if such a response is successfully corroborated by other forms of evidence. However, it is the occurrence of false identifications that has serious implications for the execution of justice. False recognitions could lead to miscarriages of justice through wrongful conviction of innocent persons. In the present study, hits and false positives were subsequently combined to obtain a more sensitive measure of recognition accuracy called A' (pronounced: A *prime*).[2] A' scores were computed for the purpose of establishing the extent to which differences in age affect face recognition performance after controlling for 'subject response bias' problems that typically occur in stimulus recognition experiments. Otherwise, hits and false positives were the primary dependent variables for the present study.

Three predictions were tested in this experiment. First, on the basis of a multi-dimensional space framework of face encoding (see Valentine, 1991a; 1991b; Valentine & Endo, 1991; and Valentine & Ferrara, 1991 for a detailed account of this model), a significant main effect of distinctiveness was expected to emerge from an analysis of hits, false positives and A' scores. Overall, participants were expected to be significantly more accurate in their recognition of distinctive faces than in their recognition of typical faces. Second, a significant main effect of age was hypothesised in which, regardless of face distinctiveness or typicality, recognition accuracy was expected to improve with increasing age. Finally, and most importantly, we predicted a significant interaction between age and face distinctiveness. The effect of distinctiveness on all the dependent variables was expected to be moderated by differences in age.

Stimuli

For use in the present experiment, 40 pairs of face photographs were taken.[3] Each pair comprised a full-front view and a three-quarter view of the same person's face. All the 40 persons whose faces were photographed were aged between 16–18 years. The faces were *ranked* from the 'most distinctive' to the''least distinctive' (i.e. typical') by 72 children and adults who did not take part in the recognition memory experiment. The procedure used to obtain the distinctiveness rankings was as follows.

1 It has been shown that face recognition performance reaches the adult level at the age of 16 years (Ellis, 1990).

2 The formula used for computing A' scores was as follows (*Source:* Rae, 1976): If h> or = f, then: $A' = [(h^2 + f^2 + 3h - f - 4fh)/4h(1 - f)]$ where: h=p(H) and f=f(FP); if h<f, then: $A' = (h - h^2 + f - f^2)/[4f(1 - f)]$.

3 The photographs were 45mm x 55mm in size (*on screen*).

All the 40 full front view faces were placed on a rectangular table measuring 2m x 1m x 1m. The photographs were arranged in rows of 10 yielding four rows on the table. There were no spaces left between the photographs. Each of the children and adults who participated in the ranking exercise was asked to stand as close as possible to the table so as to have a full view of all the faces. Each participant was then asked to examine all the photographs carefully and pick the face he/she thought would be easiest to remember if seen again. After making a choice, the selected photograph was removed from the set and given the appropriate rank value. This procedure was repeated until all the photographs had been assigned a "ranked memorability" score for each of the participants. On the basis of these rankings, mean distinctiveness values were computed for each face. Finally, the obtained means were rank-ordered to determine whether each face in the set could be classified as 'distinctive' or 'typical' using the median as the point of separation between these two conceptual categories. Sixteen of the 20 'most distinctive' and 16 of the 20 'most typical' faces were selected to serve as stimuli for this experiment. Statistical analyses performed on the raw distinctiveness data showed that the rank order values were broadly consistent across all age groups. As such, the categories of 'distinctive' and 'typical' faces generated using the procedure outlined above were deemed to be a true reflection of the perceived distinctiveness or typicality of each of the faces used in the recognition memory experiment.

In order to effectively control the time that each face was shown to each of the participants (exposure duration) and the time that elapsed between presentations (inter-stimulus interval), we decided against a 'manual' presentation of the stimuli in favour of a computerised procedure. For this reason, each of the selected faces was carefully scanned into a 486DX-100 IBM-compatible computer using a MICROGRAFIX programme linked to a Genius scanner controlled by SCANJET software. All the photographs were equalised for contrast, brightness and size using an advanced software package specifically designed for this purpose. A programme written in the C++ language was used to control face presentation times and the inter-stimulus interval, as well as to log each participant's response and response latencies.

Procedure

The experiment was conducted in rooms provided by the schools at which the study was conducted. All the 74 subjects were tested *individually*. Each subject sat in front of a personal computer at an approximate distance of 40cm from the screen. The experimenter sat next to the subject. The subject was then told that he/she was about to take part in an experiment in which his/her ability to remember faces was to be examined. Shortly after this, the 16 target faces were presented one by one on the computer screen. Each face was presented for approximately five seconds using an inter-stimulus interval of two seconds.

After all the 16 faces had been shown, subjects were informed that they were about to see 32 faces (all in three-quarter view) comprising 16 target faces (faces presented earlier) and 16 'new' faces (i.e. distracters). Target and distracter faces were shown in a quasi-random order in which the only constraint was that no more than three target or distracter faces were to appear consecutively. Subjects responded to each face by pressing on the keyboard either a key marked 'yes' (for faces they thought were targets) or one marked 'no' (for faces they thought were distracters or foils). Each participant's response to each of the 32 faces that were shown at test were logged by the computer which also collected response latency data.

Results

For each subject, the proportion of hits and of false positives were calculated. An *Fmax* test of homogeneity of variance was applied to the raw data before conducting all of the analyses reported in this article. Although the hits and false positives data did not violate the principle of homogeneity of variance for ANOVA, the A′ scores did ($F=6.01$, $p<.05$). Therefore, the raw A′ scores were subjected to an arcsin transformation before being analysed (see Winer, Brown & Michels, 1991; pp 104–105). The mean percentage hits and false positives made by the four groups of participants on recognition memory for distinctive and typical faces are shown in table 1.

Table 1:

Age Group			
6–7 year-olds			
Hits	75.00 % (sd = 6.94)	47.83 % (sd = 6.25)	61.41 % (sd = 6.59)
False positives	11.38 % (sd = 6.31)	40.44 % (sd = 6.68)	25.91 % (sd = 6.49)
9–10 year-olds			
Hits	81.25 % (sd = 6.94)	55.94 % (sd = 6.68)	68.58 % (sd = 6.81)
False positives	13.19 % (sd = 4.63)	30.94 % (sd = 4.44)	14.71 % (sd = 4.55)
12–13 year-olds			
Hits	78.81 % (sd = 4.88)	56.69 % (sd = 8.06)	68.75 % (sd = 6.47)

Separate analyses were performed on hits, false positives and A′ scores. In each analysis, the raw or transformed data were subjected to a 4 x 2 split-plot analysis of variance.

Analysis of Hits

The predicted main effect of distinctiveness was highly significant ($F(1,69)=389.38$, $p<.0001$). On the whole and disregarding age differences, participants made a significantly greater number of hits in response to distinctive faces than they made in response to typical faces. Also, the predicted main effect of age was statistically significant ($F(3,69)=21.93$, $p<.0001$). This main effect is shown in Figure 1a.

Fig. 1: Graphs showing the main effect of age and the interaction between age and distinctiveness on hits

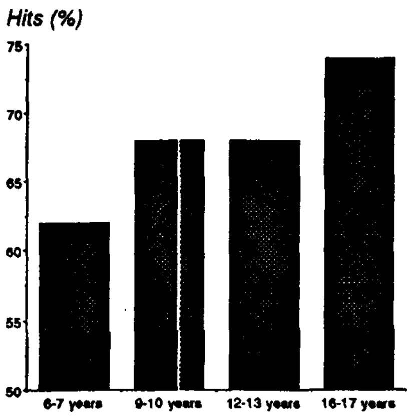


Fig 1a. Main effect of age

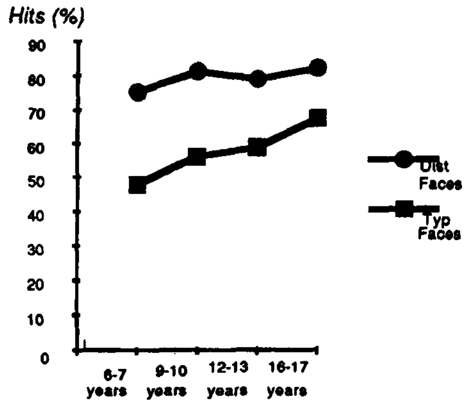


Fig1b. Age X distinctiveness interaction

It is clear from Figure 1a that the number of hits increased as a function of age although not much improvement was found between ages 9 and 13 years. Finally, and most importantly, the interaction between age and distinctiveness was significant ($F(3,69)=6.80$, $p<.001$). This interaction is shown in Figure 1b. As can be seen from Figure 1b, the main effect of distinctiveness was moderated by age. Multiple comparison tests[4] applied to the means obtained on hits showed that while age differences did not significantly affect the number of hits made by participants in response to distinctive faces ($p>.05$), differences in age significantly affected the number of hits that were made in response to typical faces ($p<.05$).

Analysis of false positives

Analysis of the false positives data revealed significant main effects of distinctiveness ($F(F1,69)=296,23$, $p<.0001$), and of age group ($F(3,69)=80.70$, $p<.0001$). Participants made a significantly greater number of false positives in response to typical faces than they made in response to distinctive faces. The main effect of age is shown in Figure 2a. Clearly, the total number of false positives (i.e. false recognitions) made declined as a function of increasing age. However, as predicted, the main effect of distinctiveness was moderated by differences in age. This was demonstrated in the highly significant interaction between distinctiveness and age ($F(3,69)=53.96$, $<.0001$) which is shown in Figure 2b. This interaction shows quite clearly that while age did not affect the number of false recognitions made in response to distinctive faces, differences in age did affect the number of false recognitions made in response to typical faces.

Fig 2: Graphs showing the main effect of age and the interaction between age and distinctiveness on false recognitions

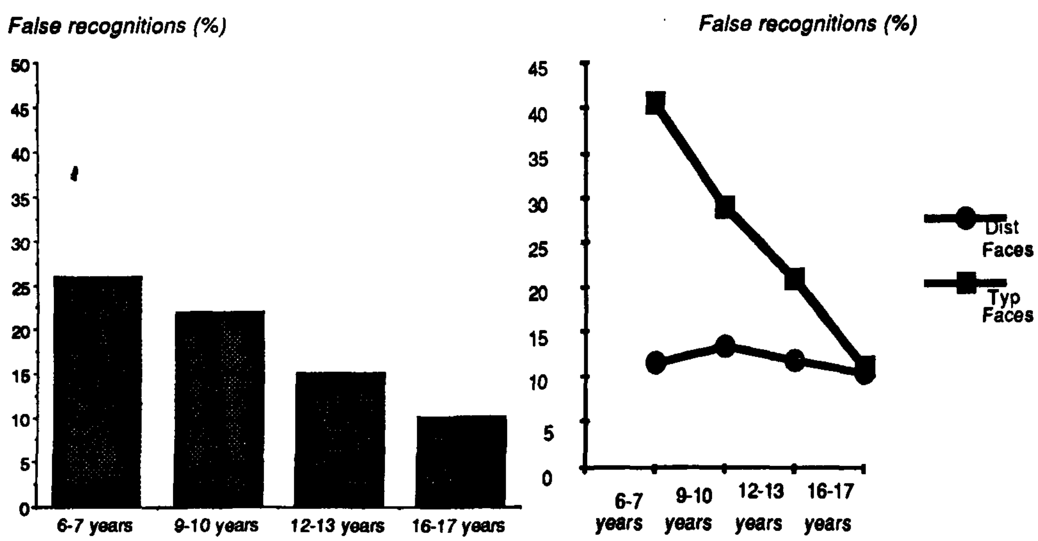


Fig 2a. Main effect of age

Fig 2b. Age X distinctiveness interaction

4 The formula used to make *all* multiple comparison tests referred to in this article was as follows:

$$t=\frac{c_j(\overline{A}_j)+c_{j'}(\overline{A}_{j'})}{\sqrt{2\,MS_{subj.w.groups}/nq}};\; df=p(n-1);\; pcrit = 0.05. \text{ (Kirk, 1969, p. 266).}$$

Discussion

A number of conclusions can be drawn from the results obtained in this experiment. First, an analysis of the proportion of hits and of false positives showed that the participants were significantly more accurate in their recognition of faces whose features and/or feature configurations made them distinctive within the set of stimuli used than they were in their recognition of typical faces. This conclusion is based on the highly significant main effects of distinctiveness that emerged from an analysis of hits and false positives data. It is likely that this finding could have resulted from less interference effects in memory between representations of distinctive faces compared to representations of typical faces. At a theoretical level, this account is consistent with the multi-dimensional space framework of face encoding proposed by Valentine (1991) and previous face encoding models such as Rhodes, Tan, Brake, & Taylor (1989)'s norm-based coding model and Valentine & Bruce (1986)'s face prototype hypothesis.

Second, the significant main effect of age obtained on hits and on false positives suggests that the ability to discriminate between individual photographs of faces seen previously is, to a large extent, dependent upon maturational factors that influence cognitive processing of faces as visual stimuli. However, it is not clear at present what the specific skills that develop with age are or whether people who differ on other psychological variables (e.g. sex, field dependence, personality traits, etc.) develop such skills at the same rate. An analysis of the A' data suggests that one of the differences between children and adults in face identification performance might be related to the issue of 'criterion setting'. As pointed out earlier, the procedure followed in computing A' scores is based on signal detection theory. According to this theory, subjects presented with a 'yes/no' identification task set different criteria for responding 'yes' to a presented stimulus. Signal detection parameters such as A' or d' (pronounced, d-*prime*) combine hits and false positives using statistical procedures that control for differences in 'response criteria' set by participants in a recognition task. An analysis of the A' scores produced F values that were far greater than were obtained from an analysis of either hits or false positives alone. It is therefore likely that due to less developed face processing skills, children tend to set a more relaxed response criterion than adults, resulting in the observed main effect of age.

Third, a statistically significant interaction between age and distinctiveness was obtained from an analysis of hits and of false positives. An examination of Figures 1b and 2b showed a clear pattern of results. When the 'to be remembered' faces were distinctive in appearance, the performance of children (even 6–7 year olds) was comparable to that of adults. However, when the target faces were typical, thus sharing several features and/or feature configurations with many other faces represented in memory, clear developmental patterns emerged from the data. Therefore, the differential levels of performance observed between children and adults on recognition of distinctive versus typical faces appears to have resulted from greater interference effects that may have been experienced by younger participants when the recognition task involved discriminating between many faces that shared similar features and/or feature configurations. Future experimental work should examine whether, compared to adults, children of different age groups use different encoding strategies when confronted with a task in which distinctive and typical faces must be recognised.

Finally, the results obtained in the present study are based on recognition of previously presented 'photographs' of faces that were seen briefly in only one view and presented in a different view at test. In real life, witnesses are often asked to identify face photographs of *people* whom they will have seen either committing a crime or near the scene of a crime. Furthermore, the witnesses will often have had the opportunity to encode the target faces from different views. On ecological validity grounds, it could therefore be argued that the results obtained in the present study may not hold if the 'to be remembered' face photograph were to be presented in a more natural setting. In addition, all the participants were tested for their recognition of target faces *soon after* they had seen them. In real life situations, witnesses are asked to identify suspects several weeks or months after a crime is committed. On the basis of these considerations, we proceeded to conduct experiment 2 using a more realistic stimulus presentation method and a reasonable delay interval between initial encounter with target faces and the face identification exercise.

EXPERIMENT 2

In this experiment, we tested a large number of children and adults on their recognition of face photographs of two research assistants who visited schools, showed a video on malaria prevention and asked the participants a few general questions before leaving the school. The time between the initial visit to the schools by the research assistants and the face identification exercise was exactly 3 months. Instead of the computerised presentation procedure used in experiment 1, photo spreads were used in the present study.

Participants

One hundred and forty four children and adults participated in the experiment. Of these, 48 were 6-7 year-olds, 48 were 9–10 year olds, and 48 were 16–18 year-olds. In each age group, equal numbers of male and female participants were included. All the participants were pupils at two schools located in one of Harare's high density suburbs.

Procedure

Two research assistants (posing as 'health workers') visited the schools and showed the selected groups of participants a video on malaria prevention. For each age group, all the participants were taken to a specially prepared room where the 'health workers' introduced themselves, their institutional affiliation, and explained the purpose of their visit (i.e. "to show a video on malaria prevention"). Participants were then taken in pairs to an adjacent room where the video was shown. Afterwards, the participants answered some questions that were asked by one of the 'health workers' before being dismissed. This procedure was followed at all the schools visited. Three months later, the same participants were asked to identify the two 'health workers' using the photo spread technique. Another research assistant who had never been to the target schools before was recruited to conduct the identification exercise. Before going to the schools, the assistant received thorough training on how the identification process was to be conducted with each participant.

First, face photographs of the two 'health workers' were taken in two views (a full front view and a three-quarter view). These photographs served as the target faces. One of the research assistants had a clearly distinctive face while the other assistant's face was typical in appearance. For each of the two target faces, 4 photographs of faces which served as distracters were obtained. Using multiple copies of all the photographs, eight photo spreads were prepared in brown albums. Five pictures were included in each of the eight albums. In four of the albums, the faces were shown in full-front view (the **A** albums) while in the remaining four albums (the **B** albums), the same faces were presented in three-quarter view. Album No. 1A contained the distinctive target face plus four distracter faces (target present condition — distinctive faces). Album No. 2A contained five distracter faces for the distinctive target face (target absent condition — distinctive faces). Album No. 3A contained the typical target face plus four typical distracter faces (target present condition — typical faces). Finally, album No. 4A contained five distracter faces for the typical target face (target absent — typical faces). Albums 1B, 2B, 3B, and 4B contained the same faces as in 1A, 2A, 3A, and 4A respectively but in three-quarter view.

Participants in each age group were randomly assigned to each of the following four experimental conditions: *condition 1*: Albums 1A & 4B; *condition 2*: Albums 2A & 3B; *condition 3*: Albums 1B & 4A; and *condition 4*: Albums 2B & 3A. Thus, each participant saw and responded to two sets of photographs. In one set, the target was present while in the other set, the target was absent. Also, in each condition, one set of photographs contained distinctive faces while the other set contained typical faces. The identification exercise was conducted in as relaxed a manner as possible. The research assistant started by greeting the participant and asking for his/her name. She then introduced herself as a social worker after which she asked the participant a few general questions aimed at relaxing the participant. At an appropriate point in the conversation, the research assistant reminded the participant about a visit to the school by two 'health workers' "... who showed you a video on malaria prevention". Once the participant confirmed that he/she recalled there being such a visit, the identification exercise started. At this point, the research assistant said the following:

> I have with me here two photo albums. I will show you the albums one by one. There are five pictures in each album. I would like you to look at each set of pictures carefully and point at the photograph which you think is that of one of the two 'health workers' who came to your school three months ago.

Each participant's choice was recorded on a pre-designed data recording sheet and subsequently entered into a data analysis programme loaded into a 486DX-100 IBM compatible computer.

Results and Discussion

Table 2 shows the percentages of participants in each age group who made false recognitions in response to each of the eight photo spreads.

Table 2: Percentages of participants who made false recognitions in response in Experiment 2

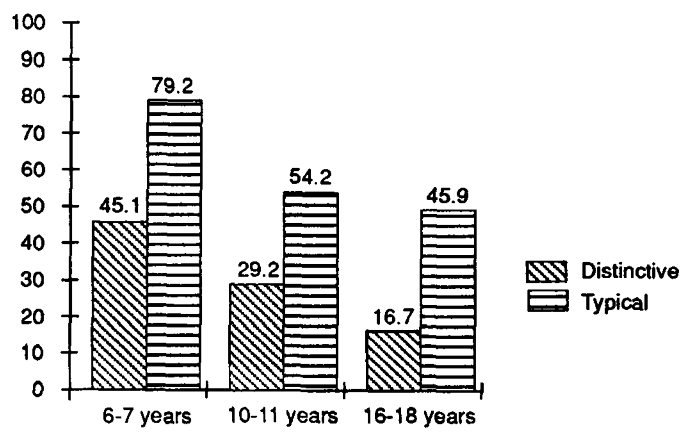
Age				
6–7 years	45.2%	100%	79.1%	100%
10-11 years	29.1%	100%	54.2%	100%

An examination of the 'target present' frequency data shows that, on the whole, a significantly greater number of participants made errors in response to the photo spreads in which the target face was typical (mean = 59.8%) than in response to the photo spreads in which the target face was distinctive (mean = 30.4%). Thus, the overall advantage for recognition of distinctive faces over typical faces that was found in experiment 1 was reliably replicated in the present experiment. It is also clear from Table 2 that disregarding face distinctiveness or typicality, the number of participants who made false recognitions declined substantially as a function of age. The percentage of participants who made false recognitions in the youngest group in the target-present condition was 62.2%. For the 10–11 year-olds, the percentage was 41.7% while for the adults, the figure dropped to 31.3%.

A striking but worrying result that is also apparent from the percentages shown in Table 2 above is that, in response to the 'target-absent' photo spreads, all the participants across all age groups made false recognitions except one adult participant who made a single correct response to a photo spread in which distinctive faces were included.

It will be recalled that in experiment 1, the decline in false recognitions observed with increasing age was restricted to recognition memory for typical faces. In that experiment, differences in age did not affect recognition memory for distinctive faces. In order to establish whether the same trend was present in the data obtained in experiment 2, frequency scores for the two target faces in the 'target present' condition were subjected to a 2 x 2 chi-square analysis. The analysis showed significant chi-square values for both the distinctive face (=6.30, df = 2; p<.04) and the typical face (=7.95, df=2; p<.01). The percentage of participants who gave incorrect responses to the distinctive and the typical target face for each age group are shown in Figure 3 below.

Fig. 3. Percentages of participants who made false recognitions in response to two photo spreads in which the target face was either distinctive or typical



It is clear from Figure 3 that the proportion of participants who made false recognitions declined with age. This was equally the case for both the typical and the distinctive target faces. In order to establish whether this decline was statistically significant, the Mantel-Haenszel test for association was applied to both sets of data. The results of this analysis showed highly significant values for the distinctive face (MH=5.47, df=1; p<.01) and for the typical face (MH=7.49, df=1; p<.006).

SUMMARY AND CONCLUSIONS

The two experiments presented in this paper clearly show that general face recognition ability and face identification skills develop and improve as a function of increasing age. The results obtained in both experiments suggest that young children aged between 6 and 7 years are particularly more likely to make false recognitions than are adults. Also, older children aged between 9 and 13 years made a significantly higher number of face recognition and identification errors than were made by adults. It is however, important to note that although adult participants made the least number of false recognitions in both experiments, their level of performance was not particularly impressive. For instance, while the mean number of false positives obtained by adults who were tested immediately after exposure to the target faces was comparatively small (10.54%), in experiment 2 where there was a three-month delay between exposure to target faces and the face identification exercise, the proportion of adults who made false identifications in the 'target present' condition was quite high (31.3%).

In Zimbabwe and many other countries, a considerable amount of time often elapses between a witness seeing a criminal and the time the witness is required to identify the suspect from a photo spread or police line-up. Thus, the results obtained in experiment 2 clearly show that after a delay of three months, even adults are significantly more likely to make face identication errors than when the identification exercise is conducted soon after exposure to a target face. For both groups of children tested in experiment 2 (i.e. 6–7 year olds and 9–10 year olds), the proportion of participants who made false recognitions was well above chance for the typical target face (66.7%) and close to chance for the distinctive target face (close to 40%) [chance = 50%]. It is therefore important for personnel in the judicial system to ensure that in cases where a particular trial rests primarily on face identification evidence, such evidence is obtained expeditiously, particularly when children are key witnesses to the crime.

Furthermore, when the target face was absent from the photo spread, both children and adults made false recognitions that were exceedingly high. Except for one adult who responded correctly once in the 'target absent' condition of distinctive faces, 100% false recognitions were made by participants in all the age groups. High percentages of false recognitions in 'target-absent' line-ups have been reported in the literature (see Malpass & Devine, 1981a for a review). One possible reason for this finding might be related to the instructions given to participants in experiment 2. Participants were not warned before hand that the target face may not be present in the photo spread. As a result, participants may have used the "best resemblance/best choice" strategy. It is possible that had participants been pre-warned, the number of false recognitions might have been lower. However, results from studies in which participants were pre-warned still showed very high levels of false recognitions by adults (e.g. Malpass & Devine obtained false recognition scores above 70%). Future research should examine whether informing 'witnesses' about the possibility of there being no target face in the photo spread or line-up would differentially reduce the number of false recognitions made by children and adults.

An important application of 'target absent' photo spreads is that of eliminating witnesses whose memory of the suspect may be unreliable. For example, in the case of John Demjanjuk, had Mrs Radiwker and her colleagues been conducting their investigations in a truly scientific and unbiased manner, they could have used photo spreads in which the target face was absent as a way of excluding witnesses who selected foils as targets (see Buchanan, 1985; Wells *et al*, 1979; Cutler, Penrod & Martens, 1987; and Wells, 1990 for more on the use of this strategy and other procedures that can be used to enhance the fairness of line-ups).

The experiments reported in this article also showed that the probability of false recognitions being made by 'witnesses' depends to a large extent on the characteristics of the target face. In particular and as one might have expected, the results obtained from both experiments showed that faces with distinctive features and/or feature configurations are less likely to be mis-identified than are faces that are 'typical' or 'average-looking' in appearance. The multi-dimensional space framework of face encoding proposed by Valentine (1991) offers a parsimonious account for this effect. According to this framework, recognition of typical faces is particularly difficult because faces which share similar features and/or feature configurations are thought to be encoded and stored 'close together' in memory. As such, when a typical face is presented to a witness, an unusually high level of memory interference occurs, hence the higher number of false recognitions observed for typical compared to distinctive faces. It must be noted here that while no member of the judicial system has control over the physical appearance of a suspect/criminal, the advantage in recognition of distinctive faces over typical faces observed in the present study has important legal implications. For instance, in constructing a line-up or preparing a photo spread, it is crucial for the police to ensure that appropriate foils (as similar as possible) are used for each target person's face in order to ensure that only evidence from reliable witnesses is gathered.

Finally, the significant interaction between age and face distinctiveness that was observed in experiment 1 strongly suggests that when the target face is distinctive, differences in age may not affect recognition performance in terms of the number of false positives made while, for typical target faces, the number of false positives declines systematically as a function of increasing age. However, in experiment 2 where the identification exercise was conducted following a delay of 3 months, age differences did not appear to have differential effects on the proportion of false identifications made by the participants in response to the distinctive and typical face photo spreads. Thus, the comparable levels of performance demonstrated by children and adults on recognition of distinctive faces in experiment 1 were absent after a delay of 3 months. For both groups of children, the proportion of participants who made false recognitions was significantly smaller for both the distinctive and the typical faces photo spreads. Future research is needed to examine this phenomenon more closely. For example, it would be useful to know at what point during a three-month delay, children's recognition of distinctive faces drops to a level that is significantly lower than that of adults.

REFERENCES

Brigham, J.C., Maas, A., Snyder, L.D. & Spaulding, K. (1982). "Accuracy of Eyewitness Identification in a Field Setting". *Journal of Personality and Social Psychology*, 42, 673-681.

Buchanan, D. R. (1985). "Enhancing Eyewitness Identification: Applied Psychology for Law Enforcement Officers". *Journal of Police Science and Administration*. 13, 4, 303–309.

Cutler, B.L., Penrod, S.D. & Martens, T. K. (1987). "Improving the Reliability of Eyewitness Identification: Putting Context". *Journal of Applied Psychology*, 72,4, 629–637.

Clifford, B.R. & Scott, J. (1978). "Individual and Situational Factors in Eyewitness Testimony". *Journal of Applied Psychology*, 63, 352-359.

Doris, J. (1995). *The Suggestibility of Children's Recollections: Implications for Eyewitness Testimony*. Washington, DC: American Psychological Association.

Ellis, H.D. (1990). "Developmental Trends in Face Recognition". *The Psychologist*, 3, 114–119.

Fisher, R. P., Geiselman, R. E., & Amador, M. (1989). "Field Test of the Cognitive Interview: Enhancing the Recollection of Actual Victims and Witnesses of Crime". *Journal of Applied Psychology*. Vol. 74, 5, 722–727.

Fruzzetti, A. E., Toland, K. Teller, S. A. & Loftus, E. F. (1992). "Memory and Eyewitness Testimony". In M. Gruneberg & P. Morris (eds.). *Aspects of Memory*. London: Routledge.

Geiselman, R.E., Fisher, R. P., MacKinnon, D.P. & Holland, H.L. (1985). "Eyewitness Memory Enhancement in the Police Interview: Cognitive Retrieval Mnemonics Versus Hypnosis". *Journal of Applied Psychology*, Vol. 70, 2, 401–412.

Green, D. M. & Swets, J.A. (1966). *Signal Detection Theory and Psycho Physics*. Wiley & Sons: New York.

Kirk, R. (1969). *Experimental Design: Procedures for the Behavioural Sciences*. California: Brooks/Cole Publishing Company.

Leippe, M. R., Wells, G.L. & Ostrom, T.M. (1978). "Crime Seriousness as a Determinant of Accuracy in Eyewitness Identification". *Journal of Applied Psychology*, Vol. 63, 3, 345–352.

Macmillan, N.A. & Creelman, C.D. (1991). *Detection Theory: A User's Guide*. Cambridge University Press: Cambridge.

Malpass, R. S. & Devine, P. (1981). "Guided Memory in Eyewitness Identification". *Journal of Applied Psychology*, Vol. 66, 3, 343–350.

Morris, V. & Morris, P. E. (1985). "The influence of Question Order on Eyewitness Accuracy". *British Journal of Psychology*. Vol. 75, 365–371.

Murdock, B.B. (1982). "Recognition Memory". In C.R. Puff (ed.). *Handbook of Research Methods in Human Memory and Cognition*. Academic Press: New York.

Rae, G. (1976). "A Non-parametric Measure of Recognition Performance". *Perceptual and Motor Skills*, 42, 98.

Rhodes, G., Tan, S. Brake, S. & Carey, S. (1987). "Expertise and Configural Coding in Face Recognition". *British Journal of Psychology*, 80, 273–283.

Shepherd, J.W. (1981). "Social Factors in Face Recognition", in G.M. Davies, H.D. Ellis & J.W Shepherd (eds), *Perceiving and Remembering Faces*. London: Academic Press.

Valentine, T. (1991a). "A Unified Account of the Effects of Distinctiveness, Inversion and Race on Face Recognition". *Quarterly Journal of Experimental Psychology*, **43a**, 161-204.

Valentine, T. (1991b). "Representation and Process in Face Recognition". In R. Watt (ed.). *Vision and Visual Dysfunction*. Macmillan.

Valentine, T. & Bruce, V. (1986). "The Effects of Distinctiveness in Recognising and Classifying Faces". *Perception*, 15, 525-535.

Valentine, T. & Endo, M. (1991). "Towards an Exemplar Model of Face Processing: The Effects of Race and Distinctiveness". *Quarterly Journal of Experimental Psychology*. 44A, 671–703.

Valentine, T. & Ferrara, A. (1991). "Typicality in Categorisation, Recognition and Identification: Evidence from Face Recognition", *British Journal of Psychology*, 82, 87-102.

Wagenaar, W.A. (1988). *Identifying Ivan — A Case Study in Legal Psychology*. Herfordshire: Harvester.

Wells, G. L. (1990). "Police Line-ups as Experiments: Social Methodology as a Framework for Properly Conducted Line-ups". *Personality and Social Psychology Bulletin*, 16, 1. 106–117.

Wells, G. L. Leippe, M. R. & Ostrom, T. T. (1979). "Guidelines for Empirically Assessing the Fairness of a Line-up". *Law and Human Behaviour*, Vol. 3, 4, 285-293.

Wells, G. L. & Turtle, J. W. (1987). "Eyewitness Testimony Research: Current and Emergent Controversies". *Canadian Journal of Behavioural Science*, 19,4, 363–388.

Winer, B. J., Brown, D. & Michels, K.M. (1991). *Statistical Principles in Experimental Design*. New York: McGraw-Hill, Inc.

Wright, D. B. (1993). "Misinformation and Warnings in Eyewitness Testimony: A New Testing Procedure to Differentiate Explanations". *Memory*, 1,2, 153–166.