Applied Cognitive Psychology (in press)
Manuscript received 4 October 2001
accepted 14 March 2002

Are police video identifications fair to African-Caribbean suspects?

Tim Valentine, Niobe Harris, Anna Colom Piera and Stephen Darling
Goldsmiths College, University of London UK

Running head: Race and video identification

Address for correspondence: Professor Tim Valentine,
Department of Psychology,
Goldsmiths College,
University of London,
New Cross
London SE14 6NW
England

email: t.valentine@gold.ac.uk
Tel: 020 7919 7871
Fax: 020 7919 7873

Acknowledgements

We gratefully acknowledge the contribution of Linda Baillie and Peter Burton of West Yorkshire Police (UK) for supplying the material required for the study.

Abstract

Analysis of lineups from criminal cases has demonstrated that video technology can produce lineups that are less biased against the suspect than live lineups, and that White suspects are less likely to be identified from a live lineup than suspects of other ethnic origins. The present study assessed the fairness of video lineups of White Europeans and of African-Caribbeans used in actual criminal cases. African–Caribbean and White European participants selected the suspect from each lineup on the basis of the original witness description of the culprit. There was no reliable difference in the fairness of video lineups as a function of the ethnic origin of the lineup members. It is concluded that, within the context of the video system studied, use of video can provide a safeguard against bias against ethnic minorities that may occur in live lineups.

Introduction.

Eyewitnesses often make mistaken identifications. An eyewitness positively identifies a volunteer, who is known to be innocent, in approximately 20% of identification attempts conducted in England (Slater, 1994; Wright & McDaid, 1996). Furthermore, mistaken identification has repeatedly been shown to be a major cause of wrongful conviction in the USA (Huff; Rattner & Sagarin, 1986; Wells, Small, Penrod, Malpass, Fulero & Brimacombe, 1998) and in England and Wales (Brandon & Davies, 1973, Devlin, 1976). Identification by more than one eyewitness is no safeguard against mistaken identification. (e.g. Wells et al., 1998; Devlin, 1976). Nevertheless, the procedures and technology used to obtain identification evidence have changed little. In many jurisdictions, a simple array of photographs (photospread) is used to obtain identification evidence (e.g., USA, Canada). In England and Wales, a live lineup (identity parade) is used. West Yorkshire Police in England have developed a computer and video system for setting up video identity parades (lineups) known as VIPER (Video Identity Parade Electronic Recording). Initial evaluation of the system has produced promising results (Valentine & Heaton, 1999; Pike, Rowlands, Towell & Kemp,1999). In this paper a comparison of fairness of VIPER video lineups for White Europeans and African- Caribbeans is reported. The introduction is structured as follows: 1) The legal context for identification of suspects in England and Wales is introduced. 2) The details of the VIPER system are described. 3) Research on the effect of the medium used for identification is briefly reviewed. 4) Previous studies of lineup fairness are discussed. 5) Previous research on VIPER is described. 6) Issues relevant to the race of suspects and witnesses are briefly reviewed.

The Legal Context in England and Wales

The current procedure for conducting identification parades in England and Wales is set out in a code of practice (1995) required by the Police and Criminal Evidence Act (1984; henceforth referred to as PACE). The code includes guidelines for the following methods of identification: an identification parade; a group identification; a video film identification; and a confrontation. A group identification involves the witness viewing the suspect amongst an informal group of people (for example in a shopping mall). A confrontation (or show-up) involves the witness viewing the suspect alone, normally in a police station. The code of practice does not allow for the use of a photospread to identify a suspect. This is a major difference in common practice between England and Wales on the one hand, and the USA and Canada on the other. Nor does the code allow for identification by presentation of a single photograph of the suspect (or photo show-up). In cases in which the identity of the suspect is not known photographs may be shown to the witness (a mug-shot search). A record of the photographs shown should be kept. If a positive identification is made, the witness concerned and any other witnesses should attend an identification parade, group or video identification, unless identification is undisputed or the suspect is otherwise eliminated from enquiries.

A preference amongst the alternative methods of identification is imposed by the code of practice. If the suspect disputes identification and consents to stand on a parade (or lineup), an identification parade must be held if it is practicable to do so. A group identification, video identification or confrontation can be conducted without the suspect's consent, for example, if the suspect refuses to co-operate with any identification procedure. However, a confrontation should not take place unless the other procedures are impracticable.

The code of practice requires that the description of the suspect as first given by a potential witness be recorded. It must be given in writing to the suspect or their solicitor before any identification attempt is carried out. The arrangements and conduct of identification attempts must be the responsibility of a police officer who is not involved with the investigation of the case ("the identification officer"). A screen (one-way mirror), which prevents the participants in the parade from seeing the witness, may not be used unless there is a representative of the suspect present or the parade is recorded on video. This caveat provides protection against the possibility of an identification officer leading the witness when out of sight of the suspect.

The identification officer is responsible for the selection of participants in the parade. It is required that "The parade shall consist of at least eight people (in addition to the suspect) who so far as possible resemble the suspect in age, height, general appearance and position in life." The suspect is asked whether they have any objection to the arrangements for the parade or any participants in it and may obtain advice from a solicitor or friend, if present. The identification officer should attempt to remove grounds for objection. If it is not practicable to do so, the officer should explain why an objection couldn’t be met. The officer will then bring in the witness to view the parade. If there are multiple witnesses, the suspect may change his or her position in the parade after each witness.

Immediately before the witness inspects the parade, the identification officer must tell the witness that the person he or she saw may or may not be on the parade. The witness must be instructed to say if she/he cannot make a positive identification, but that she/he should not make a decision before looking at each member of the parade at least twice, taking as much care and time as required. Any identification should be made by indicating the number of the person concerned. A video recording or colour photograph should be taken of any identity parade or group identification.

In the case of video identification the suspect and his/her solicitor or friend must be given an opportunity to see the complete film. Steps should be taken to remove grounds for any reasonable objection. A representative of the suspect (but not the suspect personally) may be present when a video is shown to a witness.

The VIPER video identification system.

To construct a VIPER lineup, first a video clip of the suspect is recorded under standardised conditions. The distance from the camera, the eye level, lighting and background are all controlled to match the clips in the video library. The clips consist of a head and shoulders view for a period of approximately 20 seconds. Each clip starts with a full-face view, turning to a right profile view, rotating back to a left profile view and returning to a full-face view again. The verbal description of the suspect is entered into a computer database by selecting the appropriate item on a set of menus that describe facial characteristics. Each clip of a volunteer in the database has been coded on these verbal descriptors. The best matches to the entered description are then displayed on the screen in a full-face still image. At this point the search criteria can be refined and the database searched again, if necessary. The suspect then chooses eight people who they wish to be included in the video parade. Once the foils have been selected, the images can be dragged to a timeline that will specify the order in the video lineup. The suspect can choose the position in the sequence at which his or her own face will be shown. If necessary, more than one video can be recorded with the suspect in different positions. The suspect has the option for their image to appear in different positions in the lineup if multiple witnesses are to be asked to make an identification. The videos are then electronically recorded and sealed, for use at a later date. Alternatively digitally coded video lineups can be transmitted to a computer at a remote site. When the VIPER lineup is viewed a number for each member of the lineup appears in the top left-hand corner of the screen as each face is shown.

VIPER has the potential to improve the fairness and effectiveness of identification procedures, by offering a number of important advantages. First, a large video library has been compiled which may allow the construction of a less biased lineup than is possible using a smaller selection of volunteers who are available to participate in a live lineup at a specific time and location. Second, costs can be controlled because it is not necessary to pay a volunteer every time his or her image is used in a video lineup. Furthermore, the cost of parades that have to be cancelled and re-scheduled is substantially reduced. A common cause of cancelling an identity parade is that a suspect who has been bailed or a witness fails to keep a lineup appointment. A recent study found that 5% of VIPER parades were cancelled compared to 46% of live identity parades (Pike et al., 1999). Third, video may be a less threatening means of identification for victims of violent crime or cases of witness intimidation. Use of video eliminates the need for the suspect and the witness to be present at the same location simultaneously. Witnesses attending real identification parades who admitted to being nervous were significantly less likely to identify the suspect than witnesses who did not report being nervous (Ainsworth and King, 1988).

Some research suggests that sequential presentation of faces rather than simultaneous presentation of the entire lineup reduces mistaken identification with little impact on the number of correct identifications (Cutler & Penrod, 1988; Linday & Wells, 1985; Sporer, 1993). However, Memon and Bartlett (in press) report that sequential presentation reduces the rate of choosing from a lineup (whether the choices are correct or mistaken). Any advantage of sequential presentation is believed to arise because the instructions encourage witnesses to make an absolute judgement about whether each face is the culprit. Misidentification in a simultaneous lineup can arise from witnesses using a relative judgement to pick the person who most resembles the culprit (e.g. Smith, Lindsay & Pryke, 2001). Video identification naturally yields a sequential presentation. However, the PACE code of practice requires that the witness watch the entire video lineup twice before they make a decision. This requirement of the English legal system means that VIPER lineups as used in real cases cannot be conducted in a manner required to prevent a relative judgement strategy. The same requirement to view the entire video lineup twice was used for the evaluation of VIPER lineups reported below.

The effect of test medium on identification accuracy.

A number of studies have investigated the effect of the identification test medium on the ability of 'eyewitnesses' to a staged incident to subsequently make an identification (see Cutler, Berman, Penrod & Fisher, 1994 for a review). Studies have compared the effectiveness of live lineups with video and photograph arrays. The effect of identification test medium is rather modest and inconsistent between studies. Cutler et al. (1994) conclude that "based on the available research, there is no reason to believe that live lineups, video-taped lineups or photo arrays produce substantial differences in identification performance." (p. 181). The motivation behind developing the VIPER system is to improve the quality of identification procedure through construction of lineups that are fairer to the suspect because there is a wider range of potential foils available in a video database. Cutler et al.'s conclusion that the available research suggests that video is not a less effective medium than a live parade is entirely consistent with development of VIPER. The use of video provides an effective medium to access a large database rather than providing any benefit of video per se.

Lineup fairness

The fairness of identification procedures can be compared by use of a mock witness paradigm originated by Doob and Kirshenbaum (1973). A mock witness is somebody who was not present at the crime scene. The mock witness is given the description that the witness gave, and is asked to select the person they think is most likely to be the suspect from a photograph or video lineup. The purpose of a lineup is to obtain identification evidence that provides more information than that contained in a witness' verbal description of the culprit. A mock witness simulates the worst possible scenario - a witness who has no memory of the appearance of the culprit beyond that disclosed in a verbal description, and who chooses a person from the parade who best matches that description. In a perfectly fair lineup, all of the people in the parade would fit the witness' description to the same extent. Doob and Kirshenbaum argued that the probability of the suspect being chosen in a fair lineup is at the level of chance expected if all choices are spread equally across all lineup members. However, if a parade is biased

against the suspect, because some members of the parade do not fit the witness description, the suspect may be chosen at a rate above that expected by chance alone. There is a large body of research based on the mock witness paradigm (for reviews see Wells, 1993, and Wells et al, 1998).

Very few studies have used a mock witness procedure to evaluate the fairness of lineups constructed by the police in actual criminal cases. Doob and Kirshenbaum (1973) reported an analysis of a single lineup. Brigham, Ready and Spier (1990) analysed the functional size of photo lineups and photographs of live lineups from six criminal cases. In some cases more than one witness was involved and one case used more than one lineup, therefore the fairness of 12 combinations of lineups and witness descriptions were measured. Brigham, Meissner and Wasserman (1999) analysed the fairness of photo lineups and photographs of two live lineups used in 18 criminal cases. Multiple witnesses and the use of more than one lineup yielded a total of 26 lineup fairness analyses. Wells and Bradfield (1999) reported measures of the fairness of 10 lineups used in real cases (five photo lineups and five photographs of live lineups). The cases studied by Doob and Kirshenbaum (1973), and Brigham et al. (1990;1999) were selected because the defence lawyer in the case consulted the researcher as an expert in eyewitness identification. Therefore, the cases analysed in these studies are likely to consist of a highly unrepresentative sample of cases in which the defense lawyers have grounds to suspect that the fairness of the lineups can be successfully challenged. Wells and Bradfield (1999) do not state the source of their sample.

Previous Research on VIPER

A recent study evaluated a reasonably representative[1] sample of lineups used in criminal cases (Valentine and Heaton, 1999). The stimulus material were an exhaustive set of lineups available from the police officers participating in the study, for which the necessary information was available and in which legal procedures had ended. In all cases the suspect had been convicted, although not necessarily identified in the lineup. This constraint is inevitable when working with these data from real cases, because English law prevents records of identity parades being kept if the suspect is not convicted. Valentine and Heaton evaluated the use of the VIPER computerised video identification system, using a mock witness procedure to compare the fairness of 16 video lineups with photographs of 25 live lineups. All lineups had been used to obtain identification evidence in criminal cases. They found that the video identifications were significantly less biased against the suspect than the live lineups. Twenty-five percent of mock witnesses chose the suspect from the photographs of live parades, compared to 15% from video parades. The proportion expected to choose the suspect under the assumption of equiprobabilistic choice was 11%. Furthermore, the proportion of mock witnesses who selected the suspect from video lineups was not significantly greater than that expected by chance. These results suggest that there may be a considerable advantage to be gained by using video in identification procedures.

The effect of race on identification.

People tend to be less accurate in recognising faces of people of a race different from their own (e.g. Chiroro & Valentine, 1995; Valentine, Chiroro & Dixon, 1995; see Meissner & Brigham, 2001; Sporer, 2001 a,b for reviews). Meta-analyses of experimental studies show a consistent, moderate effect of cross-race face recognition. Shapiro and Penrod (1986) report an effect size (d) of 0.53. Bothwell, Brigham and Malpass (1989) report an effect size (d) of 0.7. These effect sizes equate to a cross-race effect that accounts for 6.5 to 10% the variance in face recognition performance.

The evidence relating to cross-race identification in eyewitness identification is much less clear, possibly because the requirement of an eyewitness study usually has much less statistical power than a laboratory memory experiment (see Sporer 2001b). Lineups involving ethnic minorities may be less fair because they are constructed predominately by White police officers who are less able to select a fair lineup for ethnic minorities. White participants are less selective in the photographs they chose when constructing photograph lineups of Blacks, than they are in their construction of lineups of Whites (Brigham and Ready, 1985). Target persons in Black lineups tend to be chosen by mock witnesses more often than target persons in White lineups (Brigham et al., 1990). These two studies provide evidence that lineups of Black people may be less fair than lineups of Whites. The race of lineup members, witnesses (who provided a description) and mock-witnesses have a rather limited influence on measures of lineup fairness (Lindsay, Ross, Smith and Flanigan, 1999b). Correlations between the number of mock witness choices across different combinations of race of witness and mock witness were found to be very stable for White lineups but rather unstable for Asian and Black lineups (Lindsay et al., 1999b).

A survey of the outcome of over 1500 identification attempts involving more than 600 lineups organised by the Metropolitan Police in London found that the strongest effect was that White suspects were less likely to be identified from a lineup than suspects from ethnic minorities (Wright and McDaid, 1996). This effect is a main effect of suspect race rather than an interaction between the race of the witness and of the suspect, although it is possible that the factors are confounded if witnesses are more likely to be White. This effect of race is difficult to interpret, but it is possible that it arises because identity parades involving ethnic minorities are more likely to be biased against the suspect. One possibility is that lineups constructed for ethnic minority suspects are more likely to be biased because there are fewer volunteers available to participate in a live lineup. This problem is familiar to police forces in many parts of England. Alternatively, identity parades involving ethnic minorities may be less fair because the volunteers are likely to be selected by a white Police officer. However, as described above, under the current code of practice the suspect or his/her representative can object to any of volunteers or the line-up as a whole. This provision provides at least some protection against identification officers selecting inappropriate lineup members.

The present study.

Use of video lineups could help to reduce the problem of bias against ethnic minorities, because more foils who are similar in appearance to the suspect might potentially be available from the database than might be available to stand on a live parade. However, a higher proportion of mock witnesses selected the suspect from VIPER video lineups of African-Caribbeans than from video lineups of White Europeans (23% vs. 14%: Valentine & Heaton, 1999). Although this difference is substantial, it was not significant at the 5% level (cf. the statistically significant difference between live and video parades 25% vs. 15% in the same study). This trend suggests that the pool of African – Caribbean faces in the VIPER database may not be adequate to produce acceptable video lineups. Valentine and Heaton's study was not designed to examine the effect of race, and the race of the mock witnesses was not controlled. Therefore, the possibility that VIPER video lineups may be more biased against a suspect from an ethnic minority is worthy of further investigation.

The aim of the present research was to compare the fairness of a representative sample of video identifications from actual criminal cases involving White Europeans and African-Caribbeans[2] that had been conducted under the PACE guidelines. The experimental design included two improvements compared to Valentine and Heaton's study. First, a total of 72 mock witnesses viewed each lineup (compared to 36) and the race of the mock witnesses was balanced so that the influence of the race of the mock witnesses could be examined.

The videos evaluated were selected by the experimenters from a set of forty supplied by West Yorkshire Police. The 40 lineups were the entire set of VIPER parades available at the time of the study that met the following practical and ethical constraints. (1) Legal proceedings in all cases were complete, allowing time for any appeal. (2) English law prohibits records of identification attempts from being retained in cases that do not result in a conviction. Therefore, it was necessary that the defendant had been convicted in all cases included in the study. (3) PACE requires that the first description given by a witness is recorded. Therefore, this first description was available for use in the study. Unfortunately the race of the witness who provided the description was not included in the data available to the research team.

The cases available included a significant proportion in which the culprit was described as of 'mixed race' or 'dual ethnicity'. The sample selected for evaluation was intended to be representative of the larger pool of video identification, therefore some videos of lineups consisting of people of dual White European and African – Caribbean ethnicity were included in the sample.

The code of practice for identity parades in England and Wales requires a minimum of eight foils and one suspect to stand in a lineup. Sometimes the VIPER video lineups include more than nine people. However, it all cases the suspect appeared in the first nine positions. In order to control for the size of the lineup all videos were restricted to 9 people. This follows the procedure adopted by Valentine and Heaton (1999). This procedure was used to avoid the possibility that any differences between ethnic groups would be confounded by use of larger lineups. It is important to note that all of the lineups as evaluated complied with code of practice for the conduct of identity parades in England and Wales. Therefore, the sample is representative of the minimum requirement of the courts. Excluding additional lineup members may over-estimate but cannot under-estimate any bias against the suspect. It was our aim to establish whether the minimum requirement of the law is satisfactory.

The proportion of mock witnesses who chose the suspect served as the dependent variable. This measure was used for the following reasons. (1) It can be evaluated against that expected by chance under the assumption that the choices will be spread equally across all members of the lineup by use of binomial probability. (2) A confidence interval can be calculated for each lineup to give an estimate of the error of measurement. (3) The proportion of mock witnesses who choose a suspect is correlated with the proportion of witnesses who make mistaken identifications when asked to select a person seen in a staged incident from lineups from which the 'culprit' is absent (Lindsay, Smith & Pryke, 1999a). The correlation was not significant for other common measures of lineup bias, although it should be noted that Tredoux (1999) criticised the method used to evaluate the correlation with effective size. Furthermore, Brigham et al. (1990) concluded that the proportion of mock witnesses who chose the suspect appeared to be the most useful measure of lineup bias. For discussion of the relative merits of a number of measures of lineup fairness see Tredoux (1998, 1999), Brigham and Pfeifer (1994), Brigham, Meissner and Wasserman (1999) and Lindsay, et al. (1999a).

The proportion of mock witnesses who select the suspect are measures of lineup bias. They measure the extent to which a lineup is biased against, or in favour of, the suspect. Malpass (1981) introduced a measure known as effective size to measure the distribution of mock witness choices. When effective size is calculated, no account is taken of which member of the lineup is the suspect. Therefore, the measure is described as a measure of lineup size rather than lineup bias. Tredoux (1998) pointed out that no method exists to calculate a confidence interval for effective size of a lineup, and proposed E as an alternative measure of effective size for which a confidence interval can be calculated. Following Tredoux's recommendation, we report E as a measure of lineup size

The fairness of a lineup is a function of both the composition of the lineup and the witness description. Any differences in the completeness of the witness descriptions given for lineups of White Europeans and African-Caribbeans are a potential confounding variable in the comparison between the measures of fairness. It is possible that any difference in fairness found may be attributed to more complete witness descriptions for one set of cases. This factor may be particularly important in a study of the effects of race. White participants tend to use different features in their descriptions of faces than do Black participants. For example, white participants are more likely to

refer to hair colour and texture, whereas black participants refer to the nose and mouth more than white participants and refer to a wider range of features (Ellis, Deregowski & Shepherd, 1975; see Sporer 2001b for a review). Therefore, the number of relevant visible features mentioned in the description that could be used to select somebody from the lineup was counted and served as a covariate in the analysis, so that the variance attributable to the witness description could be partialled out in the statistical analysis. A 'relevant' visible feature is any descriptor on which any member of a lineup differs from any other lineup member. For example, if the witness described the culprit as male and all people in the lineup were male, this would not be counted as a relevant visual feature. The descriptor 'close-set eyes' would be counted because there would be some variation on this feature that could be seen in the video.

Method

Participants

There were 144 participants who acted as mock witnesses, 72 of White European origin and 72 of African-Caribbean origin. The mean age of the White European participants was 28.7 years (s.d. = 10.5); of these 52 were female (mean age = 28.6 years, s.d. = 10.83) and 20 were male (mean age = 29.0 years, s.d. = 9.9). The mean age of the African-Caribbean participants was 31.8 years (s.d. = 9.13); of these, 45 were female (mean age = 31.13 years s.d. = 8.1) and 27 were male (mean age = 33.0 years, s.d. = 10.7). The participants included members of the public. Their occupations included clerical and administrative worker, engineer, homemaker, porter and security staff, probation officer, sales representative and student. Participants were recruited from staff and students of Goldsmiths College, the Commonwealth Secretariat, and from the friends and families of previous participants.

Materials

Sixteen VIPER video lineups made by West Yorkshire Police served as stimuli. Eight of the lineups were of White European males, 6 were of African-Caribbean males and 2 were of males of White European/African-Caribbean dual ethnicity. A head and shoulders view of lineup members appeared sequentially for a period of approximately 20 seconds, starting with a full-face view, turning to a right profile view, rotating back to a left profile view and returning to a full-face view again. A number for each member of the lineup appeared in the top left-hand corner of the screen as each face was shown. Each lineup as shown to the participant consisted of a clip of the suspect and eight distractors. Fourteen of the lineups consisted of more than nine members, but only the first nine faces were shown, which in all cases included the suspect. The description of the culprit was the "first description" given by a witness that was recorded by the police when investigating the crime. In many cases more than one witness had provided a description of the culprit. In these cases only the description provided by the first witness, as listed on the case records was used.

Design and Procedure

The videos were split into two sets each comprising 4 White European, 3 African-Caribbean and 1 video lineup in which each of the members were of dual ethnicity. Each mock witness viewed only one complete set. Each set was viewed by half of the participants from each ethnic origin. The number of visual features mentioned in each description that could be useful to the mock witnesses to select the suspect was recorded. This was defined as the number of visual features that could be seen in the video but did not apply equally to all lineup members (For example, 'male' or 'White' would not be counted, but 'unshaven' or 'pale complexion' would be included.)

Witnesses were tested at a number of locations, including Goldsmiths College. A television monitor and a video cassette player were used to display the videos. Participants were informed that the study was 'double blind'. Videos were shown in a random order. The witness description of the culprit was read before each video was shown, but mock witnesses were free to ask for the description to be repeated as many times as necessary, or to enquire about specific features whilst the lineup was running. In line with the PACE guidelines, mock witnesses were advised that they were free to view the lineup as many times as they required, but that in any event, they were to view the lineup at least twice before making a decision. Witnesses were than asked to select the person who, in their judgement was most likely to be the suspect. The participants were required to make a choice (i.e., a 'don't know' response was not permitted).

Results and discussion

The mean proportion of mock witnesses who chose the suspect and the mean values of E are shown in Table 1 as a function of the race of members of the lineup and race of the mock witness[3]. The mean number of visual features mentioned in the first description of the original witnesses was 3.75 for the African-Caribbean culprits and 4.5 for the White European culprits. The frequency counts of descriptors used are shown as in Table 2.

The proportion of mock witnesses who chose the suspect is the only measure of lineup bias that was subjected to analysis for the reasons set out in the Introduction. In all of the following analyses, the two lineups of people of dual ethnicity are included in the sample of African-Caribbean lineups. We repeated all of the analyses reported here, excluding these two lineups and two White European lineups selected in order to improve the matching of stimulus set on the detail of the witness's verbal description. Excluding these lineups did not make any significant difference to the outcome of the analyses of the proportion of mock witnesses who chose the suspect. For brevity, only the analyses of the full stimulus set are reported. The mean proportion of mock

witnesses who chose the suspect across all lineups and all mock witnesses (i.e. both White European and African – Caribbean lineups and mock witnesses) was 16.9%. This result implies that in a criminal investigation approximately 17% of hypothetical witnesses who remember the description of the culprit that they gave to the police, but have no other memory of the culprit and do not reject the lineup, would be expected to select the suspect from these video lineups. Valentine and Heaton (1999) found that the suspect was chosen by 15.1% of 36 mock witness who viewed 16 video lineups produced in the same format by the same police force. Their sample comprised mainly White European and African-Caribbean lineups, but their mock witnesses were predominantly White Europeans. In contrast, the data in the present study are based on 72 mock witnesses who viewed each lineup, consisting of 36 white Europeans and 36 African – Caribbeans. Despite these differences in the size and sample of mock witnesses between the two studies, the estimates of lineup fairness are remarkably consistent.

		Mock Witnesses		
Lineup Race		White European	African-Caribbean	All mock witnesses
White European	Prop. S	0.163	0.205	0.184
		(0.062)	(0.060)	(0.060)
	E	4.449	3.973	4.366
		(0.631)	(0.472)	(0.545)
African-Caribbean	Prop. S	0.146	0.160	0.153
		(0.040)	(0.054)	(0.054)
	E	4.071	4.485	4.437
		(0.349)	(0.480)	(0.350)
All lineups	Prop. S	0.155	0.182*	0.169
		(0.036)	(0.040)	(0.040)
	E	4.260	4.278	4.402
		(0.352)	(0.332)	(0.313)

Table 1: Mean proportion of mock witnesses who selected the suspect and the mean value of E as a function of the race of lineup members and race of mock witness. Standard errors are in parentheses. * denotes significant difference from proportion choosing suspect expected by chance (0.111).

Lineup Race	gender	ethnicity	age	clothing	height	build	hair	facial hair	skin	face shape
WE	8	8	7	8	7	5	3	4	1	2
AC	8	8	7	5	5	7	6	2	3	0
Sum	16	16	14	13	12	12	9	6	4	2

Table 2: The frequency count (max. = 8) of visual descriptors included in witness descriptions as a function of the race of culprit. Information on the race of the witness is not available. Other descriptors used were voice (1 AC), accent (1 WE, 1 AC), an object being carried (1 WE), a psychological state (1 AC), a name used (1 AC) and glasses (1 WE). Key: WE = white European; AC = African – Caribbean.

If mock witnesses choose all members of a lineup equally often, the rate of false positive identifications of the suspect would be 1 in 9 or 11% (i.e. 4 out of 36 or 8 out of 72 mock witnesses). A z-test for proportions was used to determine whether the probability of choosing the suspect from the video lineups was statistically significantly different from that expected under the assumption of equal distribution of choices across all lineup members. Only one of the data points shown in Table 1 differs significantly from chance. African-Caribbean mock witnesses chose the suspect more often than expected by chance when all lineups were taken into consideration (i.e. both White European and African – Caribbean lineups). This comparison remains significant if the two lineups of people of dual ethnicity are excluded from the analysis. The proportion of White European suspects chosen by African – Caribbean mock witnesses was the greatest (0.205). However it should be noted that this mean is based on a sample of 8 rather than 16 (or 14) lineups in the comparison that did show a significant difference from chance. These comparisons show no evidence that African – Caribbean suspects were more likely than White European suspects to be selected by mock witnesses.

The effect of the race of mock witnesses and lineup members was investigated further in analyses that took the participant or the lineup as the random factor. First, the mean proportion of suspects selected by each participant was calculated separately for White European and African-Caribbean lineups. These data were subjected to a mixed ANOVA taking the race of the mock witness and the counter-balanced factor of the set of lineups as between-participant factors and the race of the lineup members as a within-participant factor. This analysis revealed a significant effect of set ($F(1, 140)=5.43$, $MSE = .036$, $p<.05$) and significant interactions of set with mock witness race ($F(1,140)=4.73$, $MSE = .036$, $p<.05$) and with lineup race ($F(1,140)=5.13$, $MSE = .029$, $p<.05$). None of these effects of are of any theoretical significance. They reflect differences in the fairness of the two arbitrary stimulus sets. It is important to note that there was no significant effect of lineup race ($F(1,140) =$

2.46, MSE = 0.29, p = 0.12), of witness race ($F(1,140) = 1.54$, MSE = 0.36, p = 0.22), nor was the interaction between the two race factors significant ($F < 1$).

The above analysis supports the null hypothesis; there was no significant effect of lineup race or its interaction with mock witness race. It can be difficult to interpret a negative result. In these circumstances the magnitude of the effect should be considered. The effect size (d) of the main effect of lineup race was 0.26 accounting for only 1.7% of the variance (based on η^2). The effect size of the main effect of mock witness race was 0.21, accounting for 1.1% of the variance and the effect size of the interaction between lineup race and mock witness race was 0.10 accounting for 0.3% of the variance. Cohen (1988) describes an effect size of d = .20 as a small effect. The main effect of lineup race was in the opposite direction to the prediction based on previous work (Valentine & Heaton, 1999; Wright & McDaid, 1996).

It should be noted that the design had sufficient power to show a significant difference between two arbitrary groupings of video lineups that were matched for the race of the lineup members. This factor had an effect size of d = 0.40 that accounted for 3.7% of the variance. (The proportion of mock witnesses who chose the suspect in set A and set B = 0.194 and. 0.142 respectively.) Meta-analyses of previous research suggest that the cross-race effect yields effect sizes in the range d = 0.53–0.7, accounting for 6.5 – 10% of the variance in face recognition experiments (Bothwell et al., 1989; Shapiro & Penrod, 1986). The sample size used in the present study gave a power of 0.85 to detect an effect of the expected size, d = 0.5, at a confidence level of 0.05. The estimates of effect size and power suggest that the experiment had adequate power to detect a cross-race effect. Therefore, the results can reasonably be interpreted as evidence that the effects of lineup race, mock witness race or their interaction were not of the magnitude predicted from previous studies of face recognition.

The analysis by participant reported above does not control for the effect of the detail of the first description provided by the witness to the crime. The mean number of visual features mentioned for the African-Caribbean lineups was 3.75 compared to 4.5 for the White European lineups. It is possible that the African-Caribbean lineups appear to be fair only because the verbal description used to measure their fairness was less detailed than that used for the White European lineups. A previous study found that the number of visual features mentioned in the verbal description was significantly related to the proportion of mock witnesses who chose the suspect (Valentine & Heaton, 1999). Therefore, we addressed this issue by conducting an analysis of covariance, as used by Valentine and Heaton, taking the lineups as the random factor and the number of visible features mentioned in the eyewitness description as the covariate. Stimulus set and lineup race were between-item factors; mock witness race was a within-item factor. There were no significant effects. The number of visual features in the witness description was not significantly related to the proportion of mock witnesses who chose the suspect (Pearson r (16) = 0.2), nor was there any effect of mock witness race, lineup race or the interaction between these factors (all F ratios < 1).

It should be noted that the lack of an interaction between the race of lineup members and the race of mock witnesses is not directly relevant to cross-race face recognition. No memory for faces was involved because the participants were not witnesses to the original crime. The procedure used in this study required participants to select a face that best fitted a verbal description. There is no a priori reason why the race of the participant or the race of the lineup should affect performance in this task. Nevertheless the current experiment does show that the samples in White European and African – Caribbean lineups showed similar levels of fairness as assessed by the mock witness procedure.

The effect of race of lineup and of mock witnesses on the size of lineup, as measured by E, was analysed in an analysis of covariance taking the lineups as the random factor. The analysis is similar to that conducted for the proportions data reported above. There were no significant effects. The outcome of this analysis and the equivalent analysis for the proportions data reported above may reflect low statistical power due to the small number of lineups that is taken as the random factor in these analyses. It is not possible to carry out the more powerful analysis taking participants as the random factor when E is the dependent variable. E is a measure of the spread of mock witnesses' choices across lineup members, so it cannot be calculated for individual participants.

In view of the implications for potential misidentification, the range as well as the mean of the measures of lineup fairness needs to be considered. Table 3 shows the proportion choosing the suspect from the least fair white European and African-Caribbean lineups as a function of mock witness race. The value of Tredoux's E for the relevant lineups is also shown. More mock witnesses of the same race as the lineup chose the suspect than did other-race mock witnesses. Across all participants the proportion who chose the suspect, 0.440 for the least fair White European lineup and 0.403 for the least fair African-Caribbean lineup, is similar to the figure reported by Valentine and Heaton (1999) for video lineups: 0.444 for the least fair African-Caribbean lineup and 0.361 for the least fair White European lineup. The proportion choosing the suspect from photographs of the least fair "live" lineup in Valentine and Heaton's study was considerably higher (0.778).

It is interesting to note that the least fair video lineups reported in Table 3 do not have extreme values of E (3.3 – 5). This suggests that, although the suspect was chosen by approximately 40% of mock witnesses, 2 – 4 other faces in the lineup attracted approximately the number of choices that would be expected by chance.

Whenever the fairness of a sample of lineups used in real cases is measured, a question of considerable applied interest is "how many of the lineups were unfair?" The definition of an unfair lineup is fraught with difficulties (Tredoux, 1998). Valentine and Heaton (1999) argued that a naive juror might reasonably expect the probability of identifying the suspect by chance in a lineup with 8 foils to be 1/9. Therefore, they suggested that the jury should be cautioned whenever this reasonable expectation is excluded by a criterion based upon the error

of the measurement of lineup fairness. We adopt the criterion that the jury should be cautioned if the entire 95% confidence interval for the proportion of mock witnesses choosing a suspect is greater than the proportion that would be expected under the assumption of equal distribution of choice (0.111in all cases reported here). Applying this criterion to the data set for all participants, 5 lineups (31% of the sample) as evaluated in this study showed a bias against the suspect that exceeded the criterion. This figure includes three White European lineups, one African-Caribbean lineup and one lineup of people of dual-ethnicity. In comparison, 48% of 'live' lineups and 31% of video lineups in Valentine and Heaton's (1999) study exceeded the same criterion for bias against the suspect. The classification of a lineup by this method is relatively stable across mock witness of different race. Of the 5 lineups that fail to meet the criterion on data from all 72 participants, 4 also fail on data drawn from either the 36 White European alone or from the 36 African – Caribbean mock witnesses alone. Only one other lineup fails the criterion based on the data from African – Caribbean mock witnesses alone. No other lineups fail the criterion based only on data from the 36 White Europeans. It should be recalled that in both the present study and Valentine and Heaton (1999) lineups were restricted to the suspect and eight volunteers. All lineups as tested complied with the minimum requires of the code of practice in England and Wales. However, the reduction in the size of some lineups may exaggerate the proportion of 'unfair' lineups as actually used in the cases.

		Mock Witnesses		
Lineup Race		White European	African-Caribbean	Mean
White European	Prop. S	0.528	0.361	0.444
	E	3.880	3.465	3.757
African-Caribbean	Prop. S	0.333	0.472	0.403
	E	4.985	3.323	4.154

Table 3: The proportion of mock witnesses who chose the suspect in the least fair White European and African – Caribbean lineup and the value of E for each lineup.

The approach that we adopt to the issue of the fairness of a lineup contrasts with that of Brigham et al. (1999) who advocate use a combination of functional size and effective size in their Overall Fairness Index. They propose a criterion that the functional size and effective size should be greater than half of the actual number of people in the lineup (nominal size). In the present study, to meet these fairness criteria both measures should be greater than 4.5 for all lineups. Brigham et al. propose the following classification to produce the Overall Fairness Index: most unfair - positive bias (i.e. against the suspect) and inadequate lineup size; unfair - positive bias only; somewhat unfair – inadequate size only; fair – no bias and adequate size. This classification system is unstable when applied to the data reported here. When choices of all 72 participants are considered 4 lineups are unfair and 2 are somewhat unfair. Data from the 36 African - Caribbean mock witnesses results in 3 lineups being classified as most unfair, 3 being classified unfair, and 5 classified as somewhat unfair. Data from the 36 White European mock witnesses results in 4 lineups being classified unfair, and 8 classified as somewhat unfair. All of these classifications are based on data from considerably more than the 18 mock witnesses that Brigham et al. recommend.

One reason Brigham et al. prefer the use of functional size over the proportion who chose the suspect is that "it does not rely on sample size" (p. S88). However, the reliability of functional size (or any measure) does depend on sample size and is adversely affected by small sample size (Tredoux, 1998). This point is demonstrated by the results of the Overall Fairness Index applied to the present data. Therefore, we strongly prefer to follow the method advocated by Tredoux (1998) by explicitly calculating the error associated with the measures of lineup bias and size and basing the interpretation on the error as well as the mean of the relevant statistic. For this reason we have employed the use of the proportion choosing the suspect and E because a confidence interval can be calculated for individual lineups for both statistics (Tredoux, 1998). Brigham et al.'s Overall Fairness Index includes a measure of lineup size. However the case for incorporating a measure of size into a fairness index is not clear, as Brigham et al. acknowledge (p. S87-88). If the choice of mock witnesses is not greater than expected by chance (taking account of the error associated with the measure), does it matter which lineup member(s) the mock witnesses chose?

Expert testimony on the quality of the procedures used to obtain eyewitness identification is rarely, if ever, admitted in English courts. Therefore, little opportunity exists for the use of mock witnesses to assess the fairness of an identification attempt. This situation is in marked contrast to that of the United States and Canada where such expert testimony is admissible, at least in some courts (see Brigham et al, 1999 for information relating to the United States). In England, the trial judge must caution the jury about the possibility of mistaken identification following guidelines established from case law (R v. Turnbull, 1976). The "Turnbull Guidelines" seek to distinguish good from poor eyewitness evidence. The judge will direct the jury to consider factors relating to the original witness situation. For example, how long the witness had the opportunity to view the culprit, whether his or her view was restricted, the lighting conditions, whether the culprit was known to the suspect. Expert testimony is admissible when the issue before a court relates to a dispute of whether a video image from CCTV is a picture of the defendant.

Conclusions

An extensive survey of the outcome of identity parades showed that White Europeans were less likely than suspects of other ethnic groups to be identified (Wright & McDaid, 1996). This evidence suggests that the methods used in constructing lineups may less fair to ethnic minorities. A possible explanation is the difficulty that police in many parts of England experience in obtaining a sufficiently large number of volunteers from ethnic minorities to take part in live identity parades. This difficulty has provided a motivation for increased use of video identification. The ability to collect a large database of video clips of volunteers should improve the fairness of lineups, especially for ethnic minorities. Valentine and Heaton (1999) showed that the VIPER video system used by West Yorkshire Police (UK) produced lineups that are less biased than live lineups. In the present study, we undertook a systematic evaluation of the fairness of White European and African-Caribbean lineups using a mock witness procedure that included participants of both races. The results indicated that the majority of the samples of video lineups for both White Europeans and African-Caribbeans were not biased against the suspects. The proportion of all mock witnesses who chose the suspect on the basis of the original witness description alone was not significantly greater than that expected by chance. However, African -Caribbean mock witnesses did chose more suspects than expected by chance when performance was analysed across all lineups. This effect appeared to be stronger for African – Caribbean mock witnesses judging White European lineups. However, it is important to note that there was no difference in the fairness of the White European and African-Caribbean lineups. The lineups from each race that were most biased against the suspect was examined because the extremes of a sample may be the cases most likely to lead to mistaken identification. There was no evidence that African-Caribbean video lineups are likely to be more biased than video lineups of White Europeans. Approximately one third of video lineups as analysed (i.e. restricted to the suspect and eight volunteers) are unfairly biased against the suspect, by the definition that the 95% confidence interval of the proportion of mock witnesses who chose the suspect excludes the proportion expected by chance. By comparison previous work, in which fairness was measured in the same way, showed that one half of 'live' lineups were unfairly biased against the suspect by this criterion. In summary, use of computerised video identification techniques is effective in reducing the occurrence of lineups that are unfairly biased for both white European and African-Caribbean suspects.

Implications for policy.

The current Police and Criminal Evidence Act (1984) code of practice restricts the use of video lineups in England and Wales. The code imposes a hierarchy in which a live identity parade is the preferred option. The second option is a group identification – in which the witness views the suspect in an informal group, for example amongst the public in a shopping mall. Video is the third option, and should be used only if an identity parade or group identification is not practicable. As a result video is used in only in a small proportion of cases. The British Home Office is reviewing the PACE code of practice at the time of writing. Our results suggest that giving video identification equal priority with a live identity parade would allow better quality of identification evidence from less biased lineups to be obtained, at a lower cost, using a procedure that may be less threatening to victims of violent crime, and reduces potential bias against ethnic minorities.

Methods used to make formal identifications differ across jurisdictions. For example photo arrays are widely used in the USA and Canada but are not permitted under English law. Trollestrup, Turtle and Yuille (1994) reported that photo arrays were used to obtain identification evidence in 90% of a sample of 170 identification attempts included in an archival study of Canadian criminal cases. The promising results obtained from the VIPER video identification system in the present study and Valentine and Heaton's (1999) study suggest that it may offer the potential to improve the quality of identification evidence in other jurisdictions. A comparison of VIPER video parades with use of photo arrays would be particularly relevant to the US and Canadian jurisdictions. A carefully controlled study could compare the relative influence of the size of the available database (of videos or photographs), an appropriate method of selecting suitable faces from the database and the influence of the media (photographs vs. video) per se. Such a study would provide valuable analysis of VIPER's performance. It would also establish whether VIPER can be a more effective tool than the use of photo arrays. Currently the only comparisons available are between VIPER and live parades. The fact that VIPER video lineups were found to be less biased than live lineups by Valentine and Heaton (1999) may be mostly attributable to the practical constraints on constructing live lineups.

Footnotes

1. The law in England and Wales requires a minimum of eight foils and one suspect to stand in a lineup. Five of the 25 photographs of live lineups and 14 of 16 videos evaluated by Valentine & Heaton (1999) originally contained more than 9 people in the lineup. The stimuli were adapted to restrict all lineups to 9 people. Extra foils were removed from the ends of the lineup or video. This procedure was used to avoid the possibility that any advantage for video could be attributed to use of larger lineups. It is important to note that all of the lineups as evaluated complied with the code of practice for the conduct of identity parades in England and Wales. Furthermore, the conclusion of that study cannot be attributed to the larger number of video lineups that were reduced from the original size, because the proportion of mock witnesses who chose the suspect was smaller in the videos than in the live lineups

2. The terms used for ethnicity refer to police codes used to describe suspects. The term 'African – Caribbean' is used rather than 'Black', because it makes explicit that it excludes people whose ethnic origin is from the Indian sub-continent, referred to as 'Asian' in the UK.

3. To facilitate comparisons with previous studies data for each lineup including the proportion choosing the suspect, functional size, effective size and Tredoux's E, together with the appropriate 95% confidence intervals and binomial probability can be downloaded from www.valentinemoore.fsnet.co.uk/trv.

References

Ainsworth, P. B., & King, E. (1988). Witnesses' perceptions of identification parades. In: M.M. Gruneberg, P. E. Morris & R. N. Sykes (eds.) *Practical aspects of memory: Current research and issues, Vol. 1*. Chicester: Wiley (pp. 66-70.)

Bothwell, R. K. Brigham, J. C. & Malpass, R. S. (1980). Cross-racial identification of faces. *Personality and Social Psychology Bulletin*, **15**, 19-25.

Brandon, R., & Davies, C. (1973*). Wrongful imprisonment: Mistaken convictions and their consequences*. London: George Allen & Unwin

Brigham, J. C., & Pfeifer, J. E. (1994). Evaluating the fairness of lineups. In: Ross, D. F., Read, J. D., & Toglia, M. P. (eds.) *Adult eyewitness testimony: Current trends and developments.* Cambridge: Cambridge University Press: (pp. 201-222)

Brigham, J. C., Meissner, C. A., & Wasserman, A. W. (1999). Applied issues in the construction and expert assessment of photo lineups. *Applied Cognitive Psychology*, **13**, S73- S92.

Brigham, J. C., Ready, D. J. (1985). Own-race bias in lineup construction. *Law and Human Behaviour*, **9**, 415 – 424.

Brigham, J. C. Ready, D. J., & Spier, S. A. (1990). Standards for evaluating the fairness of photograph lineups. *Basic and Applied Social Psychology*, **11**, 149-163.

Chiroro, P., & Valentine, T. (1995). An investigation of the contact hypothesis of the own-race bias in face recognition. *Quarterly Journal of Experimental Psychology*, **48A**, 879-894.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cutler, B. L., Berman, G. L., Penrod, S., & Fisher, R. P. (1994). Conceptual, practical and empirical issues associated with eyewitness identification test media. In: Ross, D. F., Read, J. D., & Toglia, M. P. (eds.) *Adult eyewitness testimony: Current trends and developments.* Cambridge: Cambridge University Press: (pp. 163-181).

Cutler, B. L., & Penrod, S. (1988). Improving the reliability of eyewitness identification: lineup construction and presentation. *Journal of Applied Psychology*, **73**, 281-290.

Devlin, Lord P. (1976). *Report to the Secretary of State for the Home Department on the Departmental Committee on Evidence of Identification in Criminal Cases*. London: HMSO

Doob, A. N., & Kirshenbaum, H. M. (1973). Bias in Police lineups - partial remembering. *Journal of Police Science and Administration*, **1**, 287-293.

Ellis, H. D., Deregowski, J. B. & Shepherd, J. W. (1975). Descriptions of White and Black faces by White and Black subjects. *International Journal of Psychology*, **10**, 119-123.

Her Majesty's Stationery Office (1995) *Police and Criminal Evidence Act 1984 (s.60(1)(a) and s.66) Codes of Practice.* London: HMSO Publications.

Huff, C. R., Ratner, A., & Sagarin, E. (1986). Guilty until proven innocent: Wrongful conviction and public policy. *Crime & Delinquency***, 32**, 518-544.

Lindsay, R. C. L., Smith, S. M., & Pryke, S. (1999a). Measures of lineup fairness: Do they postdict identification accuracy*? Applied Cognitive Psychology*, **13**, S93 – S107.

Lindsay, R. C. L., Ross, D. F., Smith, S. M., & Flanigan, S. (1999b). Does race influence measures of lineup fairness? *Applied Cognitive Psychology*, **13**, S109 – S119.

Lindsay, R. C. L. & Wells, G. L. (1985). Improving eyewitness identification from lineups: simultaneous versus sequential lineup presentations. *Journal of Applied Psychology*, **70**, 556-564.

Malpass, R. S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law and Human Behaviour*, **5**, 299-309.

Meissner, C. & Brigham, J. (2001). Thirty years of investigating the own-race bias in face recognition: A meta-analytic review. *Psychology, Public Policy and Law*, **7**, 3-35

Memon, A. & Bartlett, J. (in press). The effects of verbalisation on face recognition in young and older adults. *Applied Cognitive Psychology*

Pike, G. E., Rowlands, G. Towell, N., & Kemp, R. (1999). Video Identification and simultaneous vs. sequential lineups. Paper presented to the 1st Joint Conference of the Conference of the American Psychology – Law Society and the European Association of Psychology and Law, Dublin.

R v Turnbull and others (1976) 3 *All ER* 549

Shapiro, P. N. & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, **100**, 139-156.

Slater, A. (1994). *Identification parades: A scientific evaluation.* Police Research Award Scheme. London: Police Research Group, Home Office.

Smith, S. M., Lindsay, R. C. L., & Pryke, S. (2001). Postdictors of eyewitness errors: Can false identifications be diagnosed in the cross-race situation. *Psychology, Public Policy and Law*, **7**, 153-169.

Sporer, S. L. (1993). Eyewitness identification accuracy, confidence and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology*, **78**, 22-33.

Sporer, S. L. (2001a). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy and Law*, **7**, 36-97.

Sporer, S. L. (2001b). The cross-race effect: Beyond recognition of faces in the laboratory. *Psychology, Public Policy and Law*, **7**, 170-200.

Tollestrup, P. A., Turtle, J. W., & Yuille, J. C. (1994). Actual victims and witnesses to robbery and fraud: An archival analysis. In: Ross, D. F., Read, J. D. & Toglia, M. P. (eds*.) Adult eyewitness testimony: Current trends and developments*. Cambridge: Cambridge University Press: (pp. 144-160).

Tredoux, C. G. (1998). Statistical inference on measures of lineup fairness. *Law and Human Behavior*, **22**, 217-237.

Tredoux, C. G. (1999). Statistical considerations when determining measures of lineup size and lineup bias. *Applied Cognitive Psychology*, **13**, S9 – S26

Valentine, T., Chiroro, P., & Dixon, R. (1995). An account of the own-race bias and the contact hypothesis based on a 'face-space' model of face recognition. In: Valentine, T. (ed.) *Cognitive and computational aspects of face recognition.* London: Routledge (pp. 69-94).

Valentine, T. & Heaton, P. (1999). An evaluation of the fairness of police lineups and video identifications. *Applied Cognitive Psychology*, **13**, S59 – S72

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist***, 48**, 553-571.

Wells, G. L., & Bradfield, A. L. (1999). Measuring the goodness of lineups: Parameter estimation, question effects, and limits to the mock witness pardigm. *Applied Cognitive Psychology*, **13**, S27-S39.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S. Fulero, S. M., & Brinacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, **22**, 603-647.

Wright, D. B., & McDaid, A. T. (1996). Comparing system and estimator variables using data from real lineups. *Applied Cognitive Psychology*, **10**, 75-84.