

Do strict rules and moving images increase the reliability of sequential identification procedures?

Tim Valentine, Stephen Darling,
Goldsmiths College, University of London

and

Amina Memon

University of Aberdeen

Running head: Video identification procedures

Corresponding Author:

Professor Tim Valentine,
Department of Psychology,
Goldsmiths College
New Cross
London SE14 6NW
United Kingdom

Email: t.valentine@gold.ac.uk

Acknowledgements

We wish to thank the Nuffield Foundation for funding this research project (Grant no. OPD/00227/G). We gratefully acknowledge the contribution of Dirk Franssens, University of Maastricht, who assisted with data collection, Peter Burton, Head of Imaging, West Yorkshire Police and the staff at the VIPER bureau who compiled all the VIPER lineups required. The project would not have been possible without their contribution. We also thank the Home Office and police staff who contributed to the project advisory group meetings. A summary of the research reported here is described in a review paper published in Public Interest Law Reporter (Loyola University Chicago).

Abstract

Live identification procedures in England and Wales have been replaced by use of video, which provides a sequential presentation of facial images. Sequential presentation of photographs provides some protection to innocent suspects from mistaken identification when used with strict instructions designed to prevent relative judgements (Lindsay, Lea & Fulford, 1991). However, the current procedure in England and Wales is incompatible with these strict instructions. The reported research investigated whether strict instructions would enhance the reliability of identification from video. The effect of using moving rather than still video clips was also investigated. Participants witnessed a live staged incident, and attempted to identify the culprit later from police video lineups, which were run double-blind. Strict instructions produced a significantly lower rate of correct identifications in culprit present lineups, but did not significantly reduce the rate of mistaken identification in culprit absent lineups. Moving images yielded fewer mistaken identifications in culprit absent lineups.

Do strict rules and moving images increase the reliability of sequential identification procedures?

A fair and effective criminal justice system is essential for a safe, just society. Adequate safeguards against wrongful conviction promote social well-being by guarding against the risk of an innocent person suffering whilst the true offender is left free to commit further offences. Serious concerns about the role of eyewitness identification evidence in wrongful convictions have been expressed over many years. The Devlin Report (1976) was produced in response to concern about the role of mistaken identification in wrongful convictions in England and Wales. More recently the Innocence Project in the USA has resulted in more than 180 exonerations of the wrongfully convicted by analysis of DNA evidence that was not available at the original trial. Analysis of these cases has shown that mistaken eyewitness identification is the leading cause of wrongful conviction, and was a factor in the majority of cases investigated (Scheck, Neufield and Dywer, 2000; Innocence Project, n.d.).

Identification procedures conducted in England and Wales are regulated by the Police and Criminal Evidence Act 1984 (henceforth referred to as PACE). The act requires a code of practice to be drawn up for obtaining identification evidence from eyewitnesses.¹ The 1995 code stipulated that a live lineup (identity parade) must be held if it is practicable to do so. It also specified that: a lineup must comprise at least 8 foils who resemble the suspect in 'age, height and general position in life'; the officer who conducts the identification procedure should not be involved in the investigation; the witness must be cautioned that the person they saw may or may not be present and that the suspect has the right for his/her legal representative to witness the identification procedure, or for the entire procedure to be videotaped. Despite these provisions, recent

research showed that one in five witnesses who attended an operational police live lineup mistakenly identified a foil (Valentine, Pickering & Darling, 2003; Wright & McDaid, 1996).

The PACE code of practice was last revised in 2005 (Police and Criminal Evidence Act [1984] code of practice, 2005). In the latest revision the code requires that a video identification procedure should be conducted, unless there is a reason why another procedure (e.g. a live lineup) would be more appropriate. Thus the previous presumption in favour of a live lineup has been reversed in favour of video identification. Formal identification from photographs is not permitted in England and Wales in cases in which the identity of the suspect is known.²

Video identification lineups consist of 15 second moving video sequences of a head and shoulders shot of each lineup member. Initially the participant faces the camera, then looks to their left presenting a profile view to the camera, and slowly rotates their head to look to their right and presenting the opposite profile to the camera. Finally the participant looks back at the camera presenting a full-face view again. Lineups are composed of a number of these clips (at least 9) played sequentially on a television or monitor screen. Each clip is numbered consecutively. Eight of the clips are drawn from a database of foils. One video clip is of the suspect filmed in exactly the same way as the eight foils.

West Yorkshire Police developed a system for producing video lineups known as VIPER (Video Identity Parade Electronic Recording), which is in widespread use. Previous research, has shown that VIPER video lineups from real criminal cases were fairer to the suspects than conventional 'live' lineups (Valentine & Heaton, 1999). A later study found that VIPER video lineups were equally fair to white European and African – Caribbean suspects (Valentine, Harris, Colom Piera, & Darling, 2003).

Video identification has a number of important benefits compared to live lineups.

1. Use of video can dramatically reduce the delay before an identification can be organized. Live lineups were subject to long delays to enable a selection of appropriate foils to be available to stand on a lineup (Valentine et al., 2003)
2. Approximately 50% of live lineups were cancelled (for example, due to failure of a bailed suspect to attend, failure of the witness to attend, or lack of suitable volunteers). Since the introduction of video identification the proportion of procedures cancelled has fallen to around 5% (Pike, Kemp, Brace, Allen & Rowlands, 2000).
3. Availability of a large database of video clips from which to select foils can make lineups fairer to the suspect (Valentine & Heaton, 1999). VIPER currently uses a database of approximately 12,000 video clips.
4. Use of video is less threatening to victims, who no longer have to attend an identification suite where their attacker may be physically present.

The aim of the present project is to bridge the gap between the state-of-the-art theory-driven scientific literature largely based on identification from photo-spreads, and its practical application in the context of video identification systems.

Comparisons between sequential and simultaneous lineups.

The literature suggests that the reliability of eyewitness identifications can be enhanced by the use of a sequential presentation format in which photographs of faces are presented one at a time as opposed to a simultaneous viewing of all lineup members in a photo-spread (Lindsay & Wells, 1985). Sequential procedures are intended to reduce the opportunity for witnesses to compare one lineup member with another, so that they cannot select the lineup member who most resembles the person seen previously (Wells,

1984, Lindsay & Wells, 1985). There is some empirical evidence that use of a 'relative' judgment strategy is less reliable than an 'absolute' judgment strategy in which participants decide whether a single item has been seen before without an opportunity to compare different test items (Gronlund, 2004; Kneller, Memon and Stevenage, 2001). To restrict the opportunity for relative decisions the sequential procedures include a number of constraints. The lineup administrator should be blind to the identity of the suspect. The witness is shown one face at a time. They are not told how many faces will be presented, but must decide whether each face is or is not the culprit before the next face is presented (Lindsay, Lea & Fulford, 1991). Furthermore, witnesses must not be allowed a second choice. Sequential presentation reduces the likelihood of the witness making a relative judgment (Kneller et al., 2001). Sequential presentation is now recommended as best practice in some areas of the USA and Canada (Wells, Malpass, Lindsay, Fisher, Turtle & Fulero, 2000).

Stebly, Dysart, Fulero and Lindsay (2001) carried out a meta-analytic comparison of the accuracy rates in sequential and simultaneous lineup presentations based on data from 9 published and 14 unpublished papers. When present in the lineup more witnesses identified the culprit from simultaneous lineups than from sequential lineups (50% vs. 35% respectively). Incorrect rejections of the lineup (i.e. no identification of any lineup member) were less frequent from simultaneous lineups than from sequential lineups (26% vs. 46%). When the culprit was not in the lineup, there were substantially fewer correct rejections from simultaneous lineups than from sequential lineups (49% vs. 79%), and fewer incorrect identifications of a foil from sequential lineups (28% vs. 51%). In summary, sequential presentation reduces the rate of choosing from both culprit present and culprit absent lineups. Meissner, Tredoux, Parker and McLin (2005) found that sequential lineups induce a more conservative

response criterion but do not affect discrimination accuracy. The effect is to provide some protection against mistaken identification from culprit absent lineups, but at a cost to the sensitivity of the identification procedure when the culprit is in the lineup.

The meta-analysis reported by Steblay *et al.* (2001) provided a systematic assessment of the literature, nevertheless some shortcomings in the literature should be noted. There are still only a few published studies that have tested the sequential presentation procedure. A large proportion of the studies analyzed by Steblay *et al.* were unpublished and almost 70% of studies were carried out at a single research laboratory (McQuiston-Surrett, Malpass & Tredoux, 2006). The analysis highlights a clear need to replicate and extend research findings, including further investigation of the effect of sequential presentation on identification accuracy from culprit present lineups. Studies published since Steblay *et al.*'s meta-analysis have provided further evidence that a reduced rate of correct identifications from sequential culprit present lineups is a robust phenomenon (Memon & Bartlett, 2002; Memon & Gabbert, 2003a, 2003b).

The sequential presentation procedure has been developed using photographs for identification. Therefore the question arises of whether the method would be advantageous for use with video lineups. A sequential video lineup procedure has been shown to yield lower levels of mistaken identification with little cost to correct identifications in comparison to a simultaneous video procedure (Cutler and Penrod, 1988). Cutler and Penrod's sequential and simultaneous conditions both incorporated a videotaped sequential presentation of lineup members in a moving image format. These were presented on a screen flanked by still images of all the other lineup members throughout the duration of the video in the simultaneous condition, whilst in the sequential condition they were flanked only by an image of the lineup member currently in view. The presence of these flanking still images (especially in the simultaneous

condition) make the video procedures adopted by Cutler and Penrod markedly different to those used in the VIPER system.

Two possible sources exist for the benefits observed with sequential procedures over simultaneous ones: a simple benefit of presenting one face after another, or a benefit from the imposition of additional constraints on choosing (such as prohibiting witnesses from changing their decisions, not knowing the number of lineup members etc.). Either one or both of these sources may affect the propensity for making relative judgments. Lindsay *et al* (1991) report evidence that the constraints are critical to produce the desired cut in false identifications, when the stimuli used are still photographs.

Under the PACE codes which control the implementation of lineups in England and Wales, witnesses attending video identification parades must see each lineup member at least twice prior to a decision being made. Video clips of lineup members are shown one after another. Therefore, current identification practice adopts a sequential video presentation, without the adoption of strict procedural constraints. As far as we are aware, there is no evidence available as to whether the imposition of such constraints on a sequential video identification procedure has a similar benefit as to when such constraints are applied to still images.

Our first objective, therefore, was to establish whether the reliability of video identification evidence would be enhanced if a strict sequential procedure was adopted, instead of the current instructions to view the entire lineup twice before making any identification

Identification from still or moving images.

A major difference between procedures in the USA and in England and Wales is that identification in the USA is often based on a still full-face photograph, but video

identification in England and Wales requires a moving image. The available research does not provide clear evidence whether recognition of unfamiliar faces benefits from motion in the stimulus at test (O'Toole, Roark & Abdi, 2002). Eyewitnesses will usually see the face of a criminal moving during the relevant episode in real life. Therefore, the encoding specificity principle (Tulving & Thomson, 1973), leads to the prediction that viewing a moving image in a video lineup would increase the likelihood of the culprit's face being recognized, compared to a still image. When viewing a moving image there would be more cues associated with movement and a wider range of views that are common to both the encoding and the recognition episodes.

Cutler, Berman, Penrod and Fisher (1994) review a number of studies that investigated the effect of identification test medium and the richness of the cues available in the test medium. For example, Cutler, Penrod and Martens (1987) compared the effectiveness of still full-face and profile video images with a video which shows the lineup member walking and talking. The effect of test medium and of the richness of cues was surprisingly small. Cutler and Fisher (1990) found that live and video lineups produced fewer mistaken identifications of a foil than photo-spreads, but there was no effect of the test medium on the number of correct identifications. Cutler *et al.* (1994) concluded: "With respect to current practices, the conservative conclusion is that, based on available research, there is no reason to believe that live lineups, videotaped lineups or photo arrays produce substantial differences in identification performance." (p. 181). In contrast, Egan, Pittner and Goldstein (1977) reported more correct identifications from live lineups than from photo-spreads. In summary, the currently available evidence from eyewitness studies for enhanced recognition of moving images in a lineup is ambiguous. Our second objective was to determine whether use of a lineup of moving video would yield better eyewitness identification accuracy than use of a single static full-face image.

The relationship between witness confidence and identification accuracy.

An identification made by a witness who expresses a high level of confidence in court provides highly influential evidence. In the past psychologists have argued that the confidence of an eyewitness is only weakly related to their accuracy (Bothwell, Deffenbacher & Brigham, 1987 report a mean correlation of $r = .25$ across 35 studies). Recently our understanding of the accuracy – confidence relationship has become more sophisticated. Participants who experience a wide range of different viewing conditions show a moderately strong positive relationship between their confidence and the accuracy of their eyewitness identification (Lindsay, Read & Sharma, 1998). Furthermore, the relationship between confidence and accuracy is considerably stronger amongst witnesses who identify somebody from a lineup than it is amongst people who make no identification. For non-choosers the confidence-accuracy relationship is close to zero, but choosers show a moderately strong positive relationship (Fleet, Brigham & Bothwell, 1987; Sporer, 1992, 1993, Sporer, Penrod, Read & Cutler, 1995). Sporer et al. (1995) report a mean correlation of $r = .41$ for choosers only from a meta-analysis of 30 studies, but a correlation of only $r = .12$ for non-choosers. Confidence measured before an attempted identification is not as predictive of accurate identification as confidence measured immediately after the identification attempt (Cutler & Penrod, 1989). Confidence ratings taken after the incident but prior to the identification procedure have a low correlation with identification accuracy but post-identification confidence has a moderate correlation with accuracy (Bothwell et al., 1987). Cutler and Penrod (1988) reported similar results for the relationship between pre-identification confidence and post-identification confidence with accuracy following the simultaneous and sequential video identification procedures described above. Our third objective was to assess the

relationship between witness confidence and accuracy in a video identification procedure that is very widely used by British police forces. Witness confidence was measured immediately after the incident and immediately after the lineup.

The experiment reported here was designed to address the three objectives identified above, under realistic conditions. The design was informed by consultation with an advisory board that included representatives of the police and Home Office (the government department responsible for the legal code of practice for identification). The experiment involved an unexpected live staged theft. Participants gave a description of the culprit immediately after the incident. They returned to take part in an identification procedure approximately seven days later. This delay was chosen to be typical of witnesses of real crimes, following the advice of the police members of the advisory board. A video of the culprit was taken at a local, VIPER-equipped police station. The lineup was compiled by the police and the VIPER Unit following the same procedure as used in real cases.

The proportion of witnesses who correctly identify the culprit from a lineup provides a measure of the sensitivity of the procedure. The proportion of witnesses who incorrectly identify an innocent suspect from a lineup, which does not include the culprit, provides a measure of the fairness of the procedure. The diagnosticity ratio is the ratio of the former to the latter (Wells and Lindsay, 1980). Thus the diagnosticity ratio of a lineup, is a measure of the probative value of an identification of the suspect (i.e. likelihood that the suspect is the culprit). Diagnosticity is useful because it provides a single measure of the effectiveness of a lineup. However it does so at the expense of losing information regarding the sensitivity and fairness of the procedure. Therefore, the outcome of culprit present lineups, the outcome of culprit absent lineups and the diagnosticity ratio all served as dependant variables.

In relation to the effect of lineup procedure, it was predicted that the strict sequential procedure will give fewer mistaken identifications from culprit absent lineups compared to the existing procedure because the strict procedure will restrict the opportunity to make a relative judgment. Based on the experimental evidence it was predicted that under the strict sequential procedure, participants will make fewer correct identifications from culprit present lineups. However, in line with previous research (e.g. Cutler and Penrod, 1988) it was predicted that the strict procedure will yield a higher diagnosticity ratio than the existing procedure.

Based upon the encoding specificity principle, it was predicted that there would be more correct identifications from moving images than from still images. Furthermore it was predicted that witness confidence measured immediately after the identification procedure will be more strongly associated with identification accuracy than confidence measured immediately after the incident had occurred.

Method

This research was carried out under the supervision of the School of Psychology Ethics Committee at the University of Aberdeen, and in accordance with the Code of Conduct and Ethical Principles of the British Psychological Society.

Participants

Participants were 223 students at Aberdeen University who witnessed a simulated theft. Valid data were obtained from two hundred and two participants⁴, of which 52 participants were male and 150 were female. Their mean age was 21.6 years (s.d. = 5.4 years). To reduce the potential impact of idiosyncratic effects related to the individual actor used, four different actors took part in the live simulated thefts. All

were male postgraduate students of white European appearance unlikely to be familiar to participants. Actors were aged between 23 and 25 years.

Design

The experiment was conducted in two parts. Participants observed a surprise simulated theft. Later, they attempted to identify the thief from a video lineup. Mean delay between parts 1 and 2 was 8.3 days (s.d. = 2.4, min = 2, max = 16). The experiment had 3 between-participant factors with 2 levels of each factor: the image format (moving or still images), the procedure used when the witness viewed the lineup: strict (i.e. including all of the constraints recommended in Lindsay et al., 1991) or existing (i.e. as required under current PACE guidelines) and culprit presence (culprit present / culprit absent).

Scripted Theft Scenario

Participants sat where they way could see a desk with a laptop computer on it. They had been recruited to a study on health and mood, and on arrival were given a questionnaire on general health (General Health Questionnaire 12; Goldberg, 1992) and verbal instructions for completing it. The experimenter explained that he had forgotten a second part of the questionnaires and would go to collect them. Participants were told to begin the questionnaires and that the experimenter would be return before they had finished. After the experimenter then left the room a ‘thief’, approached the desk and started to undo cables from the laptop. Apparently noticing the students for the first time, the thief appeared surprised and made an excuse for removing the computer. The experimenter returned, catching the thief in the act, and accused him of attempting to steal the laptop. The actor denied any attempt to steal

and appealed to the participants asking if they had seen him do anything. The thief then left the room.

Identification Task

To create the lineups, moving digital video of each actor had been recorded in the VIPER suite at the Grampian Police headquarters in Aberdeen. For each actor, 9 members of the VIPER database were selected to use as foils. The database was searched by entering keywords of the suspect's description (e.g. white, male, 25, short black hair). The search results were thumbnail images from which foils who matched the suspect on the relevant criteria in the code of practice (age, general appearance and position in life) were selected. The process up until this point was overseen by a police identification officer, and selection of foils and recording was carried out exactly as it would in operational circumstances.

For each actor, eight lineups were created consisting of high quality, digital, colour images. A head and shoulders shot filled the full-screen. Four lineups were culprit present (CP) lineups containing the actor and 8 foils, and four were CA lineups, containing 9 foils. The CP and CA lineups were identical, except that a foil was substituted for the culprit in CA lineups. This foil was nominated as an 'innocent suspect'. Two of the CP and two of the CA lineups for each actor were composed of moving images of lineup members, whilst the remaining two CP and two CA lineups were composed of still images of lineup members. Moving images were standard VIPER presentations: 15 second long continuously moving images, in which the lineup member turned first to their left and then to their right. Still images were single digital frame captured from the moving images, showing a full-face view, visible for 15 seconds. The position of the suspect was counterbalanced so that across the entire experiment the suspect appeared in each of positions 2 – 9 approximately an equal

number of times. This counterbalancing applied across conditions, so that there was no confound between lineup position and experimental condition. The suspect never appeared in position 1. Lineup materials were identical for strict rules and existing conditions, the difference between the conditions being a procedural manipulation.

Lineups were presented using 2 monitors which displayed identical information. These pointed in different directions, so that the viewer of each could not see the other. Participants viewed lineups on a 15 inch monitor. The experimenter's monitor was obscured by a cardboard mask, which revealed only the portion of the screen showing the number of the lineup members. This arrangement allowed the experimenter to stay blind to the presence or absence of the culprit in the lineup.

Procedure

Participants attended in small groups (mean size 6.6, s.d. = 2.4, max = 14, min = 2). After the theft scenario, the experimenter immediately explained the true purpose of the experiment, and participants were asked to give a rating of their confidence that they would be able to identify the thief (post-incident confidence). They then completed two descriptions of the culprit: a free recall description followed by a cued description. This procedure was included to simulate the process of giving a verbal description to the police. Under PACE a witness who attends an identification procedure must have previously given a description which has been recorded by the police ('the first description').

In part 2, participants carried out an identification procedure. Written and verbal instructions were administered in all experimental conditions which explained the procedure and emphasised that the culprit "may or may not be present" in the lineup, and that if the witness could not recognise the culprit, they must state that they cannot make an identification. All lineups were run double-blind. Neither the

experimenter nor the witness knew whether the culprit was present in the lineup, or the position of the culprit or innocence suspect .

In the existing procedure condition, participants watched the entire VIPER parade through twice from start to finish. They were then asked if they wished to see any of the lineup members again before making their decision, and were allowed to see any number of the lineup items again, as many times as they requested, until they were able to make an identification decision. This procedure follows the current PACE code of practice.

In the strict procedure condition, participants were shown each lineup member one at a time. After seeing a lineup member, participants could make one of three selections: 'yes' – indicating a positive identification, 'no', indicating no identification, or they could request to see the same person again. Selecting 'yes' resulted in the termination of the lineup and the identification of the relevant lineup member. No further members of the lineup members were shown to the participant. If the participant selected 'no', they were shown the next person in the lineup. Once a participant had rejected a lineup member they were not allowed to select it after seeing any subsequent lineup member. Participants were not told how many people were in the lineup. The procedure continued in this way until either a selection had been made, or all 9 lineup members had been rejected, in which case, no identification was recorded. After making their decision, participants recorded their responses to the following items. 1) A rating of their confidence in the decision they had made to the lineup (post-lineup confidence). 2) An item asking whether they compared the faces with each other or just considered one face at a time when considering their response. The letter item served as a check for the use of relative judgment or absolute judgment strategies.

Results

Culprit present lineups

Culprit present lineups were administered to 105 participants. Of these, 61 % of participants made an identification, 83 % of participants who made an identification identified the culprit. Overall, 53 % of participants who saw CP lineups made a correct identification. Figure 1 shows the percentage of participants who saw a lineup of each type and made a correct identification of the culprit, an incorrect rejection of the lineup or a mistaken identification of a foil, as a function of the two independent variables. Absolute cell frequencies are also shown.

Figure 1 about here

The numbers of participants making each response were analysed by use of a hierarchical loglinear analysis. A saturated model (with 12 cells) was constructed in which the outcome of each lineup, instructions and image format were entered as variables. Components of this model were then removed using backward elimination. Probability for removal in the model was $p < .05$. The resultant parsimonious model indicated a significant main effect of instructions. Follow up χ^2 tests indicated a significant difference in the distributions of responses between the strict and existing instructions $\chi^2(2) = 9.82, p < .01, \text{Cramér's } V = .31$. Under strict instructions the proportion of witnesses who made a correct identification was 36% (19/53 witnesses) compared to 65% (34/52 witnesses) under existing instructions (odds ratio (OR) = 0.55). The proportion who incorrectly rejected the lineup was 53% (28/53 witnesses) under strict instructions, greater than the 25% (13/52 witnesses) who rejected the lineup under the existing instructions OR = 2.12. There was no significant effect of image format on the outcome of culprit present lineups, $\chi^2(2) = 1.27, p > .1, \text{Cramér's } V = .11$. The proportion of witnesses making a correct identification did not differ

significantly for moving or still images (56% vs. 46%, 27/48 vs 26/57 witnesses, OR = 1.22), and neither did the proportion who incorrectly rejected the lineup (35% vs. 42%, 17/48 vs. 24/57 witnesses, OR = 0.83).

The proportion of participants who saw a CP lineup and made an identification (whether correct or incorrect) is presented in Table 1. The data were analysed using a hierarchical loglinear analysis, conducted with the same parameters described above (except there were 8, rather than 12 cells), which indicated a single effect of lineup rules influencing choosing rate. Participants observing strict lineups made fewer selections from the lineup than did those who observed lineups with the existing rules. Follow up χ^2 tests indicated a significant difference in the distributions of responses between the strict and existing instructions $\chi^2(1) = 8.54, p < .01, \Phi = .29$. Under strict instructions 47% of witnesses made a choice (25/53 witnesses) compared to 75% (39/52 witnesses) under existing instructions (odds ratio (OR) = 0.63). There was no significant effect of image format on choosing rate, $\chi^2(1) = .49, p > .1, \Phi = .07$. The proportion of witnesses making a choice did not differ significantly for moving or still images (65% vs. 58%, 31/48 vs 33/57 witnesses, OR = 1.12).

Culprit Absent Lineups

Culprit absent lineups were presented to 97 participants. Of these, 84 % of participants viewing CA lineups correctly rejected the lineups, whilst 16 % mistakenly identified a foil. Figure 2 shows the percentage of participants who saw a lineup of each type and made a correct rejection of the lineup or a mistaken identification of a foil, as a function of the two independent variables. Absolute cell frequencies are also shown.

Figure 2 about here

The data were analysed using a hierarchical loglinear analysis, conducted with the same parameters described for CP lineups (except with 8 cells), which indicated that the only significant effect was a main effect of image format. Follow up χ^2 tests indicated that there was a difference in the distributions of correct rejections and incorrect identifications between the moving and still images, $\chi^2(1) = 5.41$, $p < .05$, $\Phi = .24$). The proportion of witnesses who made a correct rejection when viewing still images was 75% (35/47 witnesses), lower than the equivalent proportion for moving images which was 92% (46/50 witnesses, OR = 0.82). However, despite the presence of a moderate trend, the effect of lineup instruction was not statistically significant ($\chi^2(1) = 2.85$, $p < .10$, $\Phi = -.17$). The proportion of witnesses making a correct rejection did not differ significantly between strict and existing procedures (90% vs. 77%, 44/49 vs 37/48 witnesses, OR = 1.17).

None of the participants who saw a CA lineup selected the foil nominated as the innocent suspect. Only 16 participants made a mistaken identification. If distributed evenly across all 8 foils and the innocent suspect, there would be less than 2 mistaken identifications per lineup member. It is not remarkable that no witnesses would identify an innocent suspect in a fair lineup particularly when the rate of mistaken identifications is so low (Clark, 2005).

Because no innocent suspect identifications were made, choosing rate analysis for CA lineups was redundant, as choosing rates are the same as incorrect identification rates.

Diagnosticity Ratios

The diagnosticity ratio is undefined when there are no identifications of the innocent suspect (due to division by zero). Therefore, diagnosticity could not be calculated for the nominated innocent suspect. As an alternative, diagnosticity was

calculated for the worst case scenario, in which the most frequently identified foil was treated as the innocent suspect. (Note that a different foil was selected for each of the four actors because the foils differed between lineups for different actors.) Even using this approach diagnosticity could not be calculated for each cell of the experimental design because, for some actors, there were no foil identifications when strict rules were used with moving images in the lineup. The diagnosticity ratio was greater for moving image lineups (14.1) than for still image for lineups (6.8). Diagnosticity was greater for lineups conducted under the strict sequential procedure (24.0) than for lineups conducted under the existing procedure (6.3). A higher diagnosticity ratio implies that an identification of the suspect has greater probative value that the suspect is the culprit. Statistical comparisons between the ratios were carried out using Tredoux's (1998) method. There was no significant effect of either image format ($\chi^2(1) = 1.19, p > .1, \Phi = .07$) or of the procedure employed ($\chi^2(1) = 0.92, p > .1, \Phi = .06$).

Actor

To rule out confounds related to the four actors used in this study, the hierarchical loglinear analyses reported above were repeated with actor entered as an additional covariate. In neither the CP or the CA lineups actor did emerge as a significant covariate effect in interaction with the independent variables. However, in CP but not CA lineups, actor had a significant effect on outcome. Table 2 details the number of participants making each type of response, broken down by actor.

Table 2 about here

Delay

To rule out confounds related to variation in the delay between the incident and the lineup, the hierarchical loglinear analyses were also repeated with delay

entered as a covariate. In neither the CP or the CA lineups did delay emerge as a main effect or interaction component. As a further check against potential confounds, necessary because delay varied in a continuous rather than categorical fashion, it was entered into stepwise discriminant analyses. In neither CP or CA lineups was delay a significant predictor of outcome.

Relative versus Absolute Judgements

Strict procedures are intended to reduce the possibility of witnesses using a relative judgment rather than an absolute judgment. Following the identification procedure, participants were asked ‘When you were searching for the culprit, did you compare the faces with each other in your head or just concentrate on one face at a time?’ The answers to this question were analyzed to examine whether strict procedures reduced the reported use of a relative judgment strategy. These data are shown in Table 3. There was no significant difference between the number of witnesses reporting use of a relative judgment strategy as a function of the instructions used, for either culprit present lineups, $\chi^2(1) = .174$, $p > .1$, $\Phi = .04$, or culprit absent lineups, $\chi^2(1) = .066$, $p > .1$, $\Phi = .03$: the percentage of witnesses reporting relative judgment strategy use did not differ by instruction type for either culprit present (8% vs. 6%, OR = 1.36) or culprit absent (21% vs. 19%, OR = 1.11) lineups.

Significantly greater use of relative judgments was reported for culprit absent than for culprit present lineups, irrespective of whether strict or existing procedures were used, $\chi^2(1) = 8.013$, $p < .01$, $\Phi = -.20$: the percentage of witnesses reporting relative judgments for culprit present lineups was significantly lower than for culprit absent lineups (7% vs. 20%, OR = 0.35).

Table 3 about here

Confidence-accuracy relationship.

The relationship between the confidence of witnesses and the accuracy of their identification was investigated by examining the correlation separately for people who chose from a lineup and people who rejected the lineup (Sporer, 1992). For choosers, there was a significant point-biserial correlation between post-lineup confidence and lineup outcome $r_{pb}(81) = 0.387, p < .01$, and a significant correlation between post-incident confidence and lineup outcome $r_{pb}(84) = 0.257, p < .02^5$. The correlation with accuracy was not significantly higher for post-lineup confidence than it was for post-incident confidence $t(78)=0.79, p<.001$. For non-choosers, there was no significant correlation between post-lineup confidence and outcome $r_{pb}(120) = 0.07$ or between post-incident confidence and outcome $r_{pb}(123) = 0.05^5$.

Discussion

The primary aim of the research was to establish whether use of strict viewing instructions could enhance the reliability of eyewitness identification evidence from new video identification procedures utilising sequentially presented moving video images. In doing so, we also aimed to address the theoretical issue of whether the beneficial effects previously reported (Lindsay & Wells, 1985; Lindsay, Lea & Fulford, 1991; Cutler & Penrod, 1988) for sequential lineup presentation stemmed from the use of strict rules, or from the use of sequential presentations. In the current study, all presentations were sequential; hence any benefit of strict procedures must be directly attributable to the imposition of those procedures.

Strict procedures led to fewer correct identifications of the perpetrator from culprit present lineups. Similar results from a comparisons between sequential and simultaneous lineups has been reported previously in Steblay *et al.*'s (2001) meta-analysis, and in more recent studies (Memon & Bartlett, 2002; Memon & Gabbert, 2003a, 2003b). Steblay *et al.* suggested that sequential viewing instructions are less likely

to reduce the correct identification rate under conditions that closely resembled real forensic conditions. However, our results do not support this contention. The experiment reported was carefully designed to incorporate realistic conditions, under guidance from the advisory board. Consideration was given to all of the moderator variables that Steblay *et al.* identified as rendering the differences between correct identification rates for simultaneous and sequential lineups to be “small or nonexistent” (live staged events, cautionary instructions, single perpetrators, adult witnesses asked to describe the perpetrator. See Steblay *et al.*, 2001, p. 471). The adult participants in this experiment observed an unexpected, live, staged theft by a single perpetrator. They gave a written description of the perpetrator after the incident. The lineup was conducted after a forensically-relevant delay of approximately one week. The witnesses were cautioned that the culprit may or may not be present in the lineup. The video lineups were compiled from the police database and were compiled by the police following the same procedure as used in criminal cases.

When the lineups did not contain the culprit there was no significant effect of the use of strict or existing procedures on the outcome of the identification attempt. Thus a reliably reduced rate of mistaken identification from strict procedures that has been reported in previous work was not replicated. This aspect of the results was not expected. Strict procedures are believed to work by discouraging use of a relative judgement strategy (i.e. selection of the person who most resembles the culprit). One possible explanation for the lack of a significant effect of instructions may be that the use of video, presenting one face at a time as a moving image, discourages use a relative strategy regardless of the instructions given. Analysis of self-report data supports this conclusion. Only 20% of witnesses reported using a relative strategy (comparing faces to each other) in culprit absent lineups: this did not differ between the strict and existing

procedures. Note that the responses to this question suggest that concentrating on one face at a time was a more common strategy for culprit present lineups than for culprit absent lineups (93% vs. 80% respectively). These data demonstrated that the self-report methodology was sufficiently sensitive to detect different strategies when they were used. One possible cause for this difference might be that the presence of the culprit in CP lineups provides sufficiently increased overlap of encoded features between encoding and test to allow use of an absolute 'remember' judgment.

The lack of benefit of the strict procedure reported here cannot be attributed to lack of experimental power. The power of the experiment reported ($n = 202$) is comparable to similar studies reported in the literature. Moreover we took steps to ensure the forensic relevance of the experimental design that are often not found in similar published studies.

Use of moving video clips provided more reliable identification evidence procedures compared to use of a single full-face image in CA lineups, although image movement had no effect in CP lineups. The effect size of image format even in CA lineups was small ($\Phi = .24$). The wide range of views of a face seen during a moving video sequence may have provided a richer source of cues from which it was easier to exclude faces similar to the culprit. The additional viewpoints in a video clip may be more likely to reveal distinctive features of a face, which the witness does not remember and therefore is able to conclude is not the face of the culprit. Moving images enhanced the fairness of the procedure, rather than its sensitivity. This result is consistent with previous research (Cutler & Fisher, 1990). However, contrary to predictions from encoding specificity theory, there was no effect of movement on correct identifications of the target.

An identification made by a witness who expresses a high level of confidence in court provides highly influential evidence in court. In the present experiment there was a strong relationship between confidence and accuracy for witnesses who made a choice from the lineup. People who expressed the most confidence in their choice were more likely to identify the culprit than people who expressed lower confidence. There was no significant relationship between confidence and accuracy amongst witnesses who rejected the lineup. The results are consistent with previous research that found a stronger confidence-accuracy relationship for choosers than for non-choosers (Fleet, Brigham & Bothwell, 1987; Sporer, 1992, 1993, Sporer, Penrod, Read & Cutler, 1995) and extend the finding to VIPER-format sequential video presentations.

Implications for policy

This is the first study to examine in detail the potential effects of combining strict viewing procedures with sequential moving image lineups in a format used in an operational context. The practical implications of the study are clear. Strict procedures reduced the sensitivity of the identification procedure and did not significantly enhance its fairness. On the grounds of diagnosticity and fairness there is no reason to change the current practice in England and Wales - requiring witnesses to view all faces at least twice before making any decision. On the grounds of sensitivity, the existing procedure should be preferred. The existing procedure is more likely to enable a reliable witness to identify a guilty suspect present in the lineup.

The PACE code of practice requires use moving images for video identification procedures unless the suspect does not consent to the identification procedure. The current experiment showed a reliable but small advantage for moving images. There was no benefit in terms of the rate of identifications of the culprit when present in the lineup or in terms of diagnosticity, but use of moving images did reduce mistaken identification

of foils when the culprit was not in the lineup. Therefore moving images play a small role in increasing the fairness of video lineups with no cost to the sensitivity or reliability of identification evidence.

References

- Bothwell, R. K., Deffenbacher, K. A., and Brigham, J.C. (1987). Correlations of eyewitness accuracy and confidence: Optimality hypothesis revisited. Journal of Applied Psychology, 72, 691-695.
- Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. Law and Human Behavior, 29, 575-604.
- 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. Adult eyewitness testimony: Current trends and developments. Cambridge: Cambridge University Press. (pp.163-181)
- Cutler, B. L. & Fisher, R. P. (1990). Live lineups, videotaped lineups, and photoarrays. Forensic Reports, 3, 439-448.
- Cutler, B. L., Penrod, S. D. & Martens, T. K. (1987). Improving the reliability of eyewitness identifications: Putting context into context. Journal of Applied Psychology, 72, 281-290.
- Cutler, B. L. & Penrod, S. D. (1988). Improving the reliability of eyewitness identifications: Lineup construction and presentation. Journal of Applied Psychology, 73, 281-290.
- Cutler, B. L. & Penrod, S. D. (1989). Forensically-relevant moderators of the relationship between eyewitness identification accuracy and confidence. Journal of Applied Psychology, 74, 650-652.
- 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.

Egan, D., Pittner, M. & Goldstein, A. G. (1977). Eyewitness identification: Photographs v. live models. Law and Human Behavior, 1, 199-206.

Fleet, M. L., Brigham, J. C. & Bothwell, R. K. (1987). The confidence-accuracy relationship: The effects of confidence assessment and choosing. Journal of Applied Social Psychology, 17, 171-187.

Goldberg, D. (1992). General Health Questionnaire (GHQ-12). Windsor, U.K. NFER-Nelson.

Gronlund, S. D. (2004). Sequential lineups: shift in criterion or decision strategy? Journal of Applied Psychology, 89, 362-368.

Innocence project: causes and remedies of wrongful convictions (n.d.). Retrieved 21st October 2005 from <http://www.innocenceproject.com/causes/>.

Kneller, W., Memon, A. and Stevenage, S. (2001). Simultaneous and sequential lineups: decision processes of accurate and inaccurate witnesses. Applied Cognitive Psychology, 15, 659-671.

Lindsay, D. S., Read, J. D. and Sharma, K. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely Psychological Science, 9, 215-218.

Lindsay, R.C.L., Lea, J.A. & Fulford, J.A. (1991). Sequential lineup presentation: Technique matters. Journal of Applied Psychology, 76, 741-745.

Lindsay, R.C.L., & Wells, G. L. (1985). Improving eyewitness identification from lineups: Simultaneous versus sequential presentation. Journal of Applied Psychology, 66, 343-350.

McQuiston-Surrett, D. Malpass, R. S. & Tredoux, C. G. (2006). Sequential vs. simultaneous lineups: A review of methods data and theory. Psychology, Public Policy and Law, 12, 137-169.

Meissner, C. A., Tredoux, C. G., Parker, J. F. & MacLin, O. (2005).

Eyewitness decisions in simultaneous and sequential lineups: A dual process signal detection theory analysis. Memory & Cognition, 33, 783-792

Memon, A & Gabbert, F. (2003a). Unraveling the effects of a sequential lineup. Applied Cognitive Psychology, 17, 703-714

Memon, A. & Gabbert, F. (2003b). Improving the identification accuracy of senior witnesses: Do pre-lineup questions and sequential testing help? Journal of Applied Psychology, 88, 341-347

Memon, A. & Bartlett, J.C. (2002). The effects of verbalisation on face recognition. Applied Cognitive Psychology, 16, 635-650.

Pike, G., Kemp, R., Brace, N., Allen, J. & Rowlands, G. (2000). The effectiveness of video identification parades. Proceedings of the British Psychological Society, 8, 44.

Police and Criminal Evidence Act 1984 (s.60(1)(a), s.60(1) and s.66(1)) codes of practice (2005). Downloaded 18th August 2006 from:

<http://police.homeoffice.gov.uk/operational-policing/powers-pace-codes/pace-code-intro/>

O'Toole, A. J., Roark, D.A. & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. Trends in Cognitive Sciences, 6, 261-266.

Scheck, B. Neufield, P. & Dywer, J. (2000). Actual innocence. New York: Doubleday.

Sporer, S. L. (1992). Post-dicting eyewitness accuracy: confidence, decision-times and person descriptions of choosers and non-choosers. European Journal of Social Psychology, 22, 157-180.

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., Penrod, S., Read, D. & Cutler, B. L. (1995). Choosing confidence and accuracy: A meta-analysis of the confidence-accuracy relations in eyewitness identification studies. Psychological Bulletin, 118, 315-327.

Stebly, N., Dysart, J., Fulero, S. & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. Law and Human Behavior, 25, 459-473.

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

Tulving, E. & Thomson, D. (1973) Encoding specificity and retrieval processes in memory. Psychological Review, 80, 352-73

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

Valentine, T. Harris, N., Colom Piera, A. & Darling, S. (2003). Are police video identifications fair to African – Caribbean suspects? Applied Cognitive Psychology, 17, 459-476

Valentine, T., Pickering, A. and Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups Applied Cognitive Psychology, 17, 969 – 993.

Wells, G.L. (1984). The psychology of lineup identifications. Journal of Applied Social Psychology, 14, 89-103.

Wells, G.L. & Lindsay, R.C.L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. Psychological Bulletin, 88, 776-784.

Wells, G. L., Malpass, R. S., Lindsay, R.C.L., Fisher, R.P., Turtle, J. W., & Fulero, S. (2000). From the lab to the police station: A successful application of eyewitness research. American Psychologist, 55, 581-598.

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

Footnotes

1 The Police and Criminal Evidence Act (1984) does not apply in Scotland. In general Scotland has followed similar, but not identical, practice to England and Wales.

2 A witness can examine mugshot files if the identity of a suspect is unknown. PACE Code D includes separate procedures for conducting searches of mugshot files. A record of faces viewed should be kept. However, once a suspect is identified a video identification would normally be used to obtain formal identification evidence either from the same witness or from other witnesses who have not viewed the mugshots.

3. Although the updated PACE code D does not apply in Scotland, all Scottish police forces now have the VIPER system available in their area.

4. The differences in participant numbers are accounted for by 9 participants who did not return for part 2, and 5 who were excluded after attending part 2 either because they withdrew, were personally familiar with the actor in the study, or reported seeing the actor in the interval between presentation and test.

5. The difference in the degrees of freedom for different correlations occur due to missing data. A few participants omitted confidence ratings on the scales provided and their data were excluded.

Table Captions

Table 1: The percentage and absolute frequencies (in brackets) of participants in the culprit present condition who made a choice or who rejected the lineup, broken down by image format (moving video vs. a still full-face image) and instructions (strict vs. existing).

Table 2: The percentage of witnesses and number of witnesses (in parentheses) who attained each outcome for each different actor.

Table 3: The percentage of witnesses and number of witnesses (in parentheses) who answered ‘compared with each other’ or ‘one at a time’, in response to the forced choice question: ‘When you were searching for the culprit, did you compare the faces with each other in your head or just concentrate on one face at a time?’ Data are shown as a function of culprit presence and instructions.

Figure Captions

Figure 1. The percentage of participants in the culprit present condition who made correct identifications of the culprit, incorrectly rejected the lineup or mistakenly identified a foil, plotted as a function of image format (moving video vs. a still full-face image) and instructions (strict vs. existing). Data labels are absolute cell frequencies.

Figure 2. The percentage of participants in the culprit absent condition who correctly rejected the lineup or mistakenly identified a foil, plotted as a function of image format (moving video vs. a still full-face image) and instructions (strict vs. existing). Data labels are absolute cell frequencies.

	Moving		Still	
	Existing	Strict	Existing	Strict
Chose	79 (19)	50 (12)	71 (20)	45 (13)
Rejected	21 (5)	50 (12)	29 (8)	55 (16)
Lineup				

Actor	CP Lineups			CA Lineups	
	Correct Identification	Incorrect Rejection	Foil Identification	Correct Rejection	Foil Identification
1	50 (16)	31 (10)	19 (6)	82 (23)	18 (5)
2	73 (19)	23 (6)	4 (1)	84 (21)	16 (4)
3	32 (7)	64 (14)	5 (1)	85 (17)	15 (3)
4	44 (11)	44 (11)	12 (3)	83 (20)	17 (4)

		Existing Procedure	Strict Procedure	Mean
'Compared with each other'	CP	8% (4)	6% (3)	7% (7)
	CA	21% (10)	19% (9)	20% (19)
'One at a time'	CP	92% (48)	94% (50)	93% (98)
	CA	79% (37)	81% (38)	80% (75)