Pragmatic Psychology and the perils of eyewitness identification

Inaugural Professorial lecture

University of Cape Town

Colin Getty Tredoux

Department of Psychology

28 September 2011

Introduction

Ladies and gentlemen of the audience, Madame Dean, deputy Vice Chancellors, colleagues, family – and anybody else present whom I may have forgotten - I thank you for the privilege of addressing you on this occasion. I hope that I will be able to give you a glimpse tonight of what my work has been thus far.

There are two paths along which I wish to lead you. On the first of these, we will encounter an ancient peril, one which continues to pose a grave danger to us, and in particular to the administration of criminal justice. This is the peril of relying on eyewitness identification of people as evidence of identity, but in more general terms it is the peril of treating recollections from memory as fact. These are often quite simply wrong, but this is not usually apparent to the witnesses, or to those receiving their testimony. What can be done about this problem? Our itinerary on this first path tonight is mostly about what can be done: for some 30 or 40 years a number of psychologists, criminologists, and lawyers (and sometimes hybrids of these three) have addressed this problem in practical ways. Sometimes the interventions have been simple, and successful, and sometimes they have not. I have been part of that band, and tonight I will outline some of the key ideas and interventions, as well as some of the limitations.

The second, but much briefer route tonight is a path at a slightly higher elevation to the first. We will attempt to discern, from a somewhat rarefied position, what the pragmatic or practical activity of eyewitness researchers tells us about the curious discipline that is Psychology.

The problem

There is a good chance that at some point in your lives you will be victim or witness to an event and that you will later be called to give testimony about this event. There were roughly 8 000 000 instances of contact crime (i.e. involving face to face interaction) in South Africa in the last decade[1], and although individual citizens did not suffer this equally, it is clear that there is a good chance that you will be victim or witness to one at some point. The testimony of eyewitnesses has long been of considerable importance in criminal cases, and it has also long been clear that mistaken eyewitness identifications can lead to terrible travesties of justice.

Consider the case of Adolf Beck. This Norwegian émigré to London, a widely traveled man, was attempting to raise capital for a mining concession he had obtained in his homeland, when he was

[1] http://www.saps.gov.za/statistics/reports/crimestats/2010/crime_situation_sa.pdf

accosted one morning in a London street in December 1895 by one Ms Ottilie Meissonier, in the company of a police officer. She identified him as being 'Lord Wilton de Willoughby', who had seduced her into parting with various possessions, including valuable jewellery. Beck protested his innocence, but was identified in several identification parades by 12 of 18 women who had suffered similar fates. He was put on trial, and was convicted, on the basis of these multiple eyewitness identifications, as well as on the testimony of a handwriting expert. There was no English court of criminal appeal at that time, so Beck did not have the benefit of judicial review. Beck repeatedly petitioned the Home Office, protesting his innocence, but his pleadings were dismissed. He was released after nearly 7 years of imprisonment, in 1901. However, in 1904, he was arrested, and convicted, once again, of a similar offence. By a remarkable coincidence, through the combination of a cautious judge, who delayed sentencing, and a zealous pair of sisters, the real villain (one Wilhelm Meyer), was apprehended, and Beck finally released from his surreal fate.[2]

In fact, Adolph Beck never properly recovered his wits or his livelihood, and lived only four more years. It is clear from reading letters he wrote in prison that he was driven quite crazy by the events. Specifically, he developed extensive paranoid delusions, claiming that his own solicitor and friends had conspired in his unfair imprisonment. He lived just long enough, though, to see one very important consequence of his case, namely the creation of the English court of criminal appeal.[3]

It is an easy thing, perhaps trite even, to find instances where justice has miscarried. So many criminal cases are heard, hundreds of thousands per year, that there is bound to be an error at some point. Charles Rembar, an American legal scholar, once put it that "A system of justice that has no miscarriages of justice is not a workable system of justice" (Rembar, 1980, p. 95). But it is not fair to say that courts here or elsewhere have simply shrugged the problem away. This is not true. The law has pondered the problem of eyewitness identification repeatedly, and carefully: there were no fewer than three commissions into the problem of identification in England in the 20th century[4], and in several other countries. As early as 1926, 'Justice of the peace' wrote the following in the South African Law Journal:

> ... mistaken identity is the most likely and common cause of miscarriages of justice, and such miscarriages not only shock the public conscience but give rise to doubt and uneasiness as to the administration of justice (p 287).

---

[2] See Coates, 1999 for a collection of the original documents in the Beck case.

[3] Criminal Appeal Act, 1907 (U.K.), 7 Edw. VII, c. 23.

[4] Report of the Committee of Inquiry into the case of Mr. Adolf Beck (Cmd 2315), 1904; Report of the Tribunal of Inquiry on the Arrest of Major R. 0. Sheppard, DSO, RAOC (Cmd 2497), 1925; Report of the Committee on Evidence of Identification in Criminal Cases, 1976 Cmd 338 134/135, 42.

Justice Patrick Devlin, who chaired the 1976 UK commission into identification evidence that bears his name, is particularly eloquent on the subject:

> [Identification evidence] ... is exceptionally difficult to assess. It is impervious to the usual tests. The two ways of testing a witness are by the nature of his story – is it probable and coherent? – and by his demeanour – does he appear to be honest and reliable? . . . ..[In] identification evidence there is no story; the issue rests on a single piece of observation. The state of the light, the point of observation, and the distance from the object are useful if they can show that the witness must be using his imagination. But otherwise, where there is a credible and confident assertion, they are of little use in evaluating it (In Shepherd, Ellis, & Davies, 1982, p xx).

But what has the law done, in practice, to guard against the dangers that eyewitness identifications pose? A thorough account would be too lengthy to set out here, but I have tried to do so elsewhere (Tredoux & Tredoux, 1987; Tredoux, 1996; Rust & Tredoux, 1998; Tredoux & Chiroro, 2005; see also Zeffertt, Paizes, & Skeen, 2003). In brief, the basic components of the approach in South Africa have been i) to set down rules, by judicial precedent, for how police officials should conduct identification procedures; ii) to insist that an identification parade, or lineup, be conducted, and iii) to proceed at trial with considerable caution. I have looked at some length at approaches taken in other countries, particularly in the United Kingdom, and the United States, and it cannot be said that we are out of line with those approaches. In fact, some of the safeguards we have insisted on in South Africa were absent in much of the United States until the significant changes introduced during Janet Reno's term as Attorney General, in conjunction with eyewitness researchers (see Wells, Malpass, Lindsay, Fisher, Turtle, & Fulero, 2000). Two examples should make this clear: it has long been obligatory in South Africa to warn an eyewitness that a lineup might not contain the perpetrator.[5] It is also obligatory that the officer investigating a case does not arrange or conduct the lineup.[6] These important measures were not present until recently in many United States jurisdictions, and they remain contentious in many.

The law has gone some distance on this issue, but it has not gone nearly far enough. I say not 'nearly far enough' because the scale of the problem that eyewitness identification poses did not became clear until the late 1980s, and it became clear only because of a technological revolution around evidence of identity. I mean the DNA typing or profiling pioneered by Alec Jeffreys (Jeffreys, Wilson, & Thein, 1985), and its refinement by Kary Mullis (Mullis & Faloona, 1987) with polymerase chain reaction methods. As is well known, DNA typing has a very low – perhaps infinitesimally low - probability of a mistaken match, depending on how many loci are used in the matching sequence: although there is some controversy about whether the probability of a mistaken

---

[5] *R. v. Nara Sammy* 1956 (4) (SA) 629 (T)

[6] *R. v. Masemang*, 1950 (2) (SA) 488 (A)

match with present techniques is really less than 1 in 10 billion, few doubt that it is very accurate indeed. The advent of this technology made it possible to test claims of mistaken identity much more rigorously than previously possible.

Prisons are full of people who claim that they did not commit the crimes they are incarcerated for. That this is so is a stock joke among lawyers, prison wardens, and in prison movies. In the 1980s it became possible to put these claims to the test, although this was certainly not an easy thing to do – the lengths to which the first supplicants had to go to in order to have their claims tested with DNA methods was quite extraordinary[7]. The history of how this happened in the United States, starting in 1989, and continuing to the present day, is well documented (see Scheck, Neufeld, & Dwyer, 2001). The first people to be exonerated, post-conviction, on the basis of DNA tests of biological evidence were Gary Dotson and David Vasquez, who had served 10 and 4 years in prison for the crimes of aggravated kidnapping and rape, and homicide, respectively. DNA tests showed that their DNA profiles did not match those recovered from biological material collected at the crime scenes, and indeed, in both cases matched someone else. Dotson and Vasquez were exonerated in 1989, and the news of these exonerations soon spread amongst prison inmates, and criminal lawyers, legal aid centres, and human rights organizations. Within a short period of time, multiple cases were being re-opened, and DNA tests were being conducted when biological material from the cases was still available. The exonerations stacked up, year after year, in increasing numbers: One in 1990, two in 1991, five in 1992, seven in 1994, thirteen in 1996, sixteen in 2000, eighteen in 2003, and eighteen in 2010. A total of 273 people have been granted post-conviction DNA exonerations in the USA since 1989. Of these 273, 17 had been sentenced to death before DNA led to their release. The average time spent in prison was 13 years.

I must mention here the signal contribution that 'The Innocence Project' has played in this astonishing cascade of exonerations. Barry Scheck, and Peter Neufeld, two American lawyers, founded the Innocence Project – essentially a non-profit law clinic - at Yeshiva University in the US, in 1992. It has been a driving force behind the exonerations, as well as the drive to reform criminal justice in the US, to prevent, or reduce, these specific kinds of injustice. Their work has spawned a network of innocence projects all over the world, and there is indeed a South African chapter in Johannesburg, at WITS university, known as the WITS Justice Project.

---

[7] See for instance the accounts of Kirk Bloodsworth, and Ronald Cotton, both of whom were released after more than a decade in prison through DNA testing (Junkin, 2004; Thompson-Cannino, Cotton, & Torneo, 2009) – it took several years in each case to get the DNA testing completed, and recognized by legal authorities.

The scale of the problem is vast. We still don't have a good idea of just how vast, as the 273 post-conviction DNA exonerations is almost certainly an underestimate of the number who have been convicted falsely: these 273 were released on the basis of DNA testing, but DNA tests can only be used where biological evidence is available. In the majority of cases, this is not available – there is no biological evidence, or the evidence has been destroyed, or it cannot be located. We simply don't know how many there are, but there is little doubt that it is a large number.

On a more chilling note, since 17 of the 273 prisoners who were exonerated were on 'Death Row', it seems likely that some of the people executed in the US since 1976 – more than 1000 in total - were innocent. We know that at least four executions in the United Kingdom prior to 1964 were in error.[8]

What of South Africa? There are several cases, but the one of most relevance is probably that of the so-called Eikenhof Three. These three men were tried, and two sentenced to death - later commuted to life imprisonment - for the terrible murder in 1993 of members of the Mitchley and Silberbauer families, in Eikenhof, Johannesburg. The men – Sipho Bholo, Titi Ndweni, and Sipho Gavin, ANC operatives – were identified by an eyewitness whose motor vehicle was hijacked, and used as the base from which to shoot the Mitchley family, as they drove to school. The three men confessed to this crime while in police custody, although they declared during the trial that the confessions had been extracted under duress. They went to prison protesting their innocence. While they were in prison, a member of the organization APLA (Azanian People's Liberation Army), Phila Dolo, applied for amnesty for the event in question, testifying that he had ordered three APLA operatives to carry out an attack on a school bus, and for various reasons this had failed, and the operatives had instead shot up the Mitchley family car. Several aspects of the prosecution and conviction of the three ANC operatives were problematic, apart from the questionable confessions. The details are too complex to set out here, but they include suppression of potentially exculpatory evidence by the police and prosecution, the alleged bribing of an alibi witness to renounce the alibi, and ballistics evidence linking a weapon used at the scene of the Eikenhof crime to Mr Phila Dolo. The Appeal Court heard an application in 1999 to re-open the case, and granted it, setting aside the convictions and sentences, and remitting the case to the trial court. The National Director of Public Prosecutions declined to prosecute the three men again, and they were finally released on 9 November 1999.

[8] Timothy Evans, Mahmood Mattan, George Kelly, and Derek Bentley have all been exonerated or pardoned; see http://en.wikipedia.org/wiki/Wrongful_execution.

What are the factors that led to the false conviction and imprisonment of this large number of people? There are several, and indeed in a much earlier, landmark study of such cases Edwin Borchard identified them very succinctly:

> The causes of the error are, in the main, mistaken identification, circumstantial evidence (from which erroneous inferences are drawn), or perjury, or some combination of these factors (Borchard, 1932, p viii).

Analysis of the 273 cases of DNA exoneration in the United States shows the pattern set out in Table 1. It is clear that eyewitness misidentification played a major role in these cases, but what is perhaps not so clear is that eyewitness misidentifications were often accompanied by false confessions, perjured evidence from informants, or mistaken forensic science.

Factor	Percentage of cases
Eyewitness misidentification	75
Invalid forensic science	50
False confessions and incriminating statements	25
Informants	19

Table 1. Factors implicated in the false conviction of 273 defendants (*Innocence Project*, 2011)

It is clear from this accumulation of evidence that there is a significant problem with eyewitness identification, and that the many attempts to safeguard it have not succeeded. What more can we do about the problem?

Psychology is a discipline that likes to meddle in the business of the world, and it is no surprise that it has addressed the problem of eyewitness identification, over a fairly long period of time. Early studies were conducted in the late 19th and early 20th centuries by Alfred Binet, Hugo Münsterberg, William Stern, and – this may surprise you - Hendrik Verwoerd (Binet, 1900; Münsterberg, 1908; Stern, 1910; Verwoerd, 1929). This early work was not sustained, though, and was taken up in earnest again only in the 1970s. The reasons for this renewed interest are unclear, but see some work by Siegfried Sporer (Sporer, 1982, 2008) for an historical overview.

It is not hard to see why psychologists should be interested in failures of eyewitness identification. These failures imply failures of perception, or memory, or perhaps show a proneness to suggestion – all questions that are of natural interest to cognitive and social psychologists.

How best should we tackle the problem? The approach that has prevailed is in my view characteristic of psychology as a discipline, both in what it has managed to achieve, and in its limitations. The approach has been thoroughly pragmatic, and I can summarise it for you very

simply as 'study something you can do something about'. To see the full implications of this, we need to step back a bit.

The approach that suggests itself at first is to rely on what is known as the 'two step' model of science. This is what – I suggest – most of us believe about the relationship between basic and applied science, or between basic and applied research, or between science and technology. Figure 1 is a typical diagrammatic representation of this model (it is taken from a Philosophy of Science course you can enroll for on the internet).

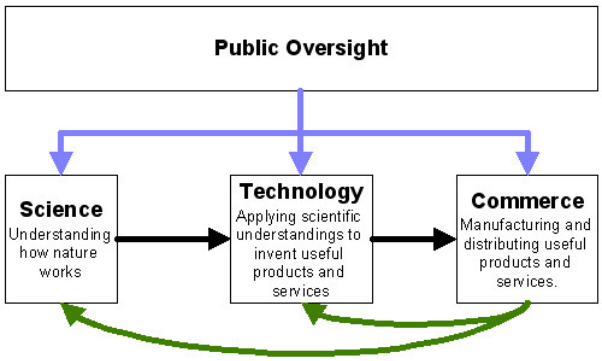

Figure 1. A typical representation of the 'two step' model of science and technology

This model suggests that we need to start in the laboratory, with basic research, aimed at recovering universal understanding of how phenomena work, before we can move to applications of the research. We ought to be able to move from laboratory research on memory and perception, for instance, to technologies that would assist us to assess an eyewitness' memory, or to improve it.

Jonathan Potter (1982) characterized this directional view of the relationship between science and technology, as the 'ideology of science': that is to say, the view has inscribed in it the assumption that science is the source of many things that make our lives better, more progressive, and healthier. Vaccines, sanitary living conditions, and Teflon frying pans, to take just a few accomplishments, presumably stemmed from basic scientific research, and have improved our lives, and this presumably warrants diverting more public money back to basic science.

But this model is questionable. Don Ihde (2010) for instance, has mounted cogent arguments that the relationship is just the other way round: "…both historically and ontologically technology precedes science" (p 117). Kurt Danziger (1990) demonstrates this very ably for the case of work on intelligence in Psychology. The first work on intelligence in Psychology was predominantly practical. Tests of intelligence were in use in France as early as 1905, and in the United States by

1910, which is long before formal theories of intelligence were mooted. More to the point, a close look at the activities of applied researchers, in multiple fields, shows a nearly complete divergence from basic research: their work is published in different outlets, citation analyses show that they cite other applied research articles more than 90% of the time[9], they typically devise new analytic methods, and as was certainly clear in the case of the intelligence researchers, develop new conceptualizations of the things they study.

I have written elsewhere about the two step model, and its problems, specifically in relation to work on memory and eyewitness identification (Tredoux, 1998a). In this area, perhaps quite unusually, many researchers have taken a conscious decision to depart from the two step model. Gary Wells' 1978 article (Wells, 1978) has been the most influential in this respect: he argued explicitly that the concepts and theories of basic research on human memory are of little use, as their domain of utility for this particular application is unclear, and will probably always be unclear. Wells made the very useful distinction between what he called 'estimator' and 'system' variables. In the case of the former, he means characteristics of the witness, or indeed any aspect of the event, that cannot be controlled in actual criminal cases. For instance, we could ask what impact the seriousness of a crime has on the ability to remember important details of the crime, but the futility of such an approach is quickly demonstrated. In particular, all we could do with this knowledge – and it would be no trivial matter to acquire it – is to give a probabilistic estimate of its effect in a particular case, and the margin of error would likely be so high as to render it of little value. Consider now a different class of characteristics, the so-called system variables in Wells' dichotomy. These are aspects of the criminal justice system that can directly be controlled or manipulated. How many people should stand in the lineup? What instruction should be given to the witness? These are the things we should investigate, argued Wells. We can dismantle the way eyewitness identifications are presently secured, and test possible modifications, using the procedural and analytic methods of experimental and social psychology. The trick is to study something you can do something about.

It has become clear to me over more than twenty years of work in this field that there is a limit to this approach, and I will argue this point later in the talk.

[9] For example, research on the U.S. weapons industry shows that 91% of innovations in the technology originated from inside the technology itself, and only 9% from basic scientific research (Potter, 1982).

What has Pragmatic Psychology done about things?

A great deal of empirical research has been conducted, and much written in Psychology journals about eyewitness identification – so much, indeed, that I have to be highly selective in what I present to you.

Let us start at a point which is near the beginning of the chain of events that ensues once a crime has been committed. If there is a witness to the sequence of events, the police will want to interview that witness and get as much relevant information as they can. How should they do this?

Edward Geiselman of UCLA, and Ron Fisher of Florida International University, have thought about this a great deal, and have worked with police for over two decades now, assessing their current methods, and working on alternative approaches. Their answer is to use an interviewing procedure known as the 'cognitive interview' (see Fisher, McCauley, & Geiselman, 1994; Fisher, 1995; Geiselman, Fisher, Mackinnon, & Holland, 1986; Fisher & Schreiber, 2007).

The cognitive interview has two key ideas embedded in it. The first is to use interviewing techniques that have been tested and refined in clinical, counseling, and research settings. The methods Fisher and Geiselman found police using were those you typically find in untrained interviewers: they interrupt witnesses a lot, and inappropriately, allow very little free response, instead of asking open-ended questions they tend to ask short, overly direct, leading questions, or complex questions to which there could be multiple answers. There is typically little effort to establish rapport or effective social dynamics, and in particular a failure to clarify the role of the eyewitness. In brief, the idea behind Geiselman and Fisher's cognitive interview is to use a set of basic, but best practice techniques when interviewing witnesses. These include, very briefly, i) developing rapport with witnesses, ii) promoting active witness participation in the interview, iii) promoting extensive, detailed responses, iv) getting responses that are code-compatible – if the information required is inherently non-verbal the interviewer should attempt to get this information in that form - for instance, if the spatial positioning of objects in a room is important, the witness should give this information by placing the objects, or substitutes for the objects, in a spatial model of the room.

The second key idea is to recognize the cognitively demanding nature of the task that an eyewitness must complete. There are several things that can be done to assist, many of which are intended to provide cues for the retrieval of information from memory. One is to reinstate the context under which the information was encoded. The idea that simply putting someone back in the encoding

context may provide relevant cues is one with a fine pedigree: Endel Tulving did a lot of the primary research in the 1960s and 1970s (e.g. Tulving & Thomson, 1973), and coined the term 'encoding specificity' to capture the idea of the empirical connectedness of memories and the context in which they are encoded. Figure 2 reports data from Godden and Baddeley's (Godden & Baddeley, 1975) famous demonstration of how scuba divers remember more information that they have encoded underwater when they are again underwater, as opposed to on dry land, and is a striking image of this connectedness.

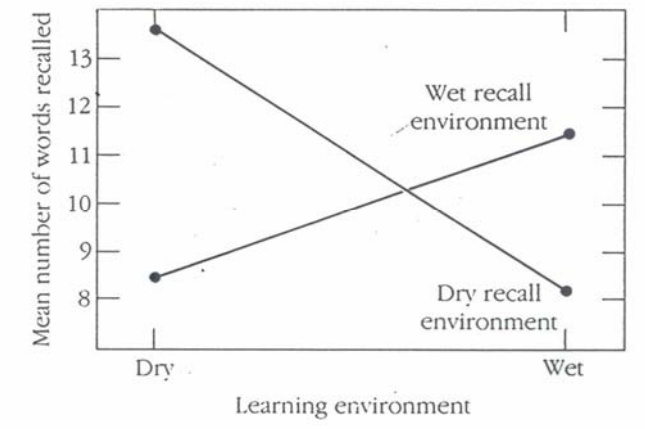

Figure 2. Context effects on recall for scuba divers (Godden & Baddeley, 1975)

Reinstating physical context is not always possible, and so researchers have explored reinstating the mental context of the event in interview, that is asking witnesses to reconstruct their affective, physiological, and cognitive states at the time of the event (see Malpass and Devine, 1981).

Secondly, interviewers can vary the kinds of retrieval cues that witnesses are given. Typically, they are trying to recollect details from a complex event, and it helps to encourage a variety of cues: witnesses are asked to re-tell their account in reverse chronology, or from a different spatial position to that which they occupied, or in terms of the sensory aspects of the event, or the semantic properties of the event.

Particularly important, given what we have learned about memory since Bartlett's pioneering studies at Cambridge in the 1930s (Bartlett, 1932), is to minimize constructive recall. This is the tendency to infer what an event should be rather than recollect it verbatim. If I ask what you ate for breakfast last Monday morning, the chances are that you will correctly answer that question, but not because you recall eating breakfast, but probably through inferring what the answer should be – if you eat Muesli every morning, with milk, for instance, that is what you will answer, but not because you remember eating it last Monday. We do this more or less automatically, and a lot of what we believe to be a veridical re-telling of our experiences is really a combination of what we in fact

recollect, and the inferences we make about what happened. When we interview eyewitnesses we want a veridical account, as far as is possible.

There have been a number of empirical tests of the cognitive interview, in a variety of settings, with students, with special populations, in laboratories, and in real contexts. A typical laboratory procedure would be to expose a group of participant witnesses to an event (live or filmed), and then to randomly assign them to a control interview procedure (a standard interview), or to a cognitive interview. The results from these studies are clear, and consistent: the cognitive interview yields more information than the control interview– approximately 34%, although this not an easy thing to quantify – of roughly the same accuracy.

However, what the cognitive interview does not do is enhance person identification – it does not enhance the accuracy with which perpetrators are identified. This is a significant failure, but one should not overlook the value to police of eliciting more accurate information with an altered interview protocol.

An element of the police interview of eyewitnesses is to get information about the appearance of the perpetrator or perpetrators. This can be elicited to some extent through verbal or written description, but faces are notoriously difficult to describe accurately. Police have used alternate means to get this information, perhaps most well known in this respect is the use of identikit portraits. I will return to the question of the utility of these later in the talk.

Improving the lineup

In many countries, a central component of the way we obtain identification evidence from eyewitnesses is to ask them to identify the perpetrator, if he (or she) is present, from a collection of people who are reasonably similar in appearance. The origin of this procedure, known as an identification parade, or a lineup, is not clear: the Metropolitan Police appear to have been using it as early as 1860 (Devlin, 1976), but there is some documentary evidence that a similar procedure may have been used, for different purposes, in Ancient Rome (Wright & Graham, 2005). A lineup seems to be intended to elicit independent evidence of identity, but also to test the hypothesis that the suspect is the perpetrator. It is seen as an important safeguard against the possibility of a mistaken identification. In the South African case of *R v Mputing* 1960 (1) SA 787, (T) the presiding Judge held (at 787) that "…where there is uncertainty as to the identity of the person who has committed an offence, an identification parade ought to be held".

Another reason for conducting lineups appears to be to replace more suggestive methods of personal identification, such as in-court or dock identifications, and direct confrontations – in each of these cases the suspect is presented individually to the witness, and the witness is asked to confirm that he or she is the perpetrator.

In almost all of the 273 exoneration cases we discussed earlier witnesses were in fact shown a lineup, and made incorrect identifications, with grave consequences, so we know that the lineup is a flawed safeguard. What to do about the lineup has been a central question for legal psychologists. The first attempt to do something substantive was in the early 1970s, when Tony Doob and Herschel Kirshenbaum published an interesting attempt to measure the fairness of police lineups (Doob & Kirshenbaum, 1973). Their technique involved giving a number of so-called mock witnesses a rudimentary description of the police suspect, and a photograph of the police lineup, and asking them to identify the suspect from the mere description. (Mock witnesses are people who have not seen the original perpetrator, or the suspect, or indeed the lineup, on any prior occasion). Several variations on this technique have been used: the mock witnesses might be shown the lineup without any description of the suspect, or might be given a consensual description of the suspect. In all cases, the idea is very simple: mock witnesses should not be able to identify the suspect, and if they are, that means the lineup is unduly suggestive. Doob & Kirshenbaum determined the proportion of mock witnesses who were able to guess the identity of the suspect, and treated this as an estimate of 'lineup bias'. Simple statistical procedures can be used to decide if this proportion is greater than one would expect from chance, namely $1/k$, where k is the number of lineup members, and to construct a statistical confidence interval around the estimate of bias. Doob and Kirshenbaum evaluated a police lineup using the measure of lineup bias, which they found in that particular case to be highly biased, and had their assessment accepted by a Canadian court.

A lineup should not be unduly suggestive, and part of what this means is that there should be a sufficient number of foils in the lineup who are plausible alternatives to the suspect – that is, who bear a sufficient degree of resemblance to the suspect. Roy Malpass pointed out in 1981 (Malpass, 1981) that the measure of bias did not necessarily index this, and proposed a measure known as 'effective size'. Although the lineup might have a nominal size of six, the five foils might not be plausible alternatives to the suspect, and the measure of lineup bias would not capture this. For instance, consider the arrays shown in Table 2, which represent choosing frequencies for 120 mock witnesses viewing a six person lineup, in each of four lineups:

Lineup	1	3	3	*Suspect*	5	6	Bias
A	20	20	20	*20*	20	20	$\frac{1}{6}$
B	0	0	0	*120*	0	0	1
C	80	0	0	*40*	0	0	$\frac{1}{3}$
D	40	0	0	*40*	0	40	$\frac{1}{3}$

Table 2. Distributions of mock witness choices in four hypothetical lineups

It is clear from visual inspection that lineups C and D have quite different patterns of choice frequencies, and that they do not contain an equal number of plausible alternatives. Malpass (1981) proposed a formula for reflecting the number of plausible alternatives, and showed how this could also be used to decide on the suitability of individual foils when constructing lineups. In other words, he gave the police something they could use proactively to create better lineups – police are very sensitive to what they often see as a critical and negative academic audience: when academics get involved, criticism will surely follow shortly afterwards!

One of my first involvements in the field of eyewitness research was to look carefully at the measures previous researchers had proposed, and to suggest improvements. In particular, Malpass' measure of effective size is a descriptive measure, but what we really needed was a measure that would allow us to draw statistical or inferential conclusions about the number of plausible alternatives. To that end, I modified a measure Alan Agresti had proposed for a different purpose in 1975, namely to measure the variation in an array of frequencies, and showed how it could be used to measure the so-called 'effective size' of a lineup (Tredoux, 1998b; Tredoux, 1999).

These measures of lineup fairness – bias, and size – are now widely used in research laboratories across the world, and expert witnesses in US courts have frequently used them to assess police lineups.

The mock witness technique is very simple, but it is really an indirect way of asking whether the lineup has satisfied a basic requirement, which is that the foils should bear a sufficient resemblance to the suspect. There are also some significant limitations to the mock witness technique: foremost is that it is a simulation of what an eyewitness with an impoverished memory will do, but we have no way of knowing what level to simulate. In practical terms, what description should we give the mock witnesses? If we use a very detailed description, the mock witnesses are very likely to identify the suspect, and if we use a brief description the mock witnesses are unlikely to identify the

suspect if the lineup meets minimal standards of suspect-foil similarity. In other words, lineup fairness and the eyewitness' encoding strength are confounded (Tredoux, Parker, & Nunez, 2007).

What would be useful is a direct measure of suspect to foil similarity. In some parts of the world, including South Africa, we insist on corporeal lineups – suspects and foils must be presented in person – but this is increasingly rare. In the UK witnesses are shown short video clips of three angles of each lineup member's face, and in the USA police mostly use photospread lineups, consisting of head and shoulder photographs of lineup members. If we are looking for a direct measure of suspect to foil similarity then it is facial similarity we need to pursue.

I first thought about this when I was starting as a graduate student in 1986, when I chanced upon an article in the Scientific American column 'computer recreations', which described a computer program for caricaturing faces, by Susan Brennan, at MIT. Brennan proposed a theory of face caricature in her MSc thesis (Brennan, 1982), later published in *Leonardo* (Brennan, 1985), which suggested that face shapes can be represented in multidimensional space – for instance, a set of co-ordinates in two dimensions, where each co-ordinate pair is set to a particular facial landmark. Once you have a representation in a multidimensional space, you can potentially measure the distance between the face and any other point, for instance the origin, or the 'average face'. She claimed that the theory of caricature which is implicit in the practice of caricature is that faces become more recognizable for their individuality when they are shifted away from the norm. She draws explicitly on Ernst Gombrich here:

> It is not really the perception of likeness for which we are originally programmed, but the noticing of unlikeness, the departure from the norm which stands out and sticks in the mind.(Gombrich, Hochberg, & Black, 1973, p 172)

I was particularly fascinated by the notion that the distance between two different faces could be navigated, as in the examples Brennan gives of the caricature and anti-caricature of the faces of celebrities. Figure 3, taken from her 1985 article, shows this:

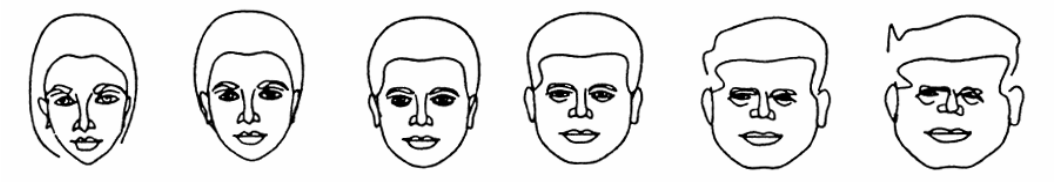

Figure 3. Caricatures and anti-caricatures of Elizabeth Taylor and John F. Kennedy : Susan Brennan's 1982 caricature generator

The idea or concept of a multidimensional face space has embedded in it the seed for an overall measure of facial similarity. At about the same time that I was first pondering this possibility Tim Valentine was demonstrating a model which presumes that faces are represented spatially in mind or brain, and that this model could account for some of the well known phenomena in face recognition research, but particularly the difficulty we have in recognizing inverted faces or faces from a different ethnic group to our own (Valentine, 1991). The problem with his model was that the dimensions structuring this space were unspecified. It struck me that they had to be functions of face images, of the percepts the space was intending to represent.

At about this point (1987 – 1992), several people in different fields – computer science, and cognitive psychology in particular – were applying image analysis methods to faces, for a number of different purposes. One group – Larry Sirovich and Robert Kirby, at Brown University (Kirby & Sirovich, 1990; Sirovich & Kirby, 1987) – was exploring ways of making face images less bandwidth intensive, in the search then for what seemed a promising technology, the videophone. They applied a statistical method well known to people who work with quantitative data, and which is known as principal component analysis (PCA) to psychologists and psychometrists. What they were able to show is that a face could be represented as a set of coefficients on an underlying set of basis vectors - things they called 'eigenfaces'. Figure 4 shows the first 8 eigenfaces returned from a PCA of 280 South African faces, with widely varying attributes, and Figure 5 shows the reconstruction of 10 faces from the original set from increasing numbers of eigenfaces.

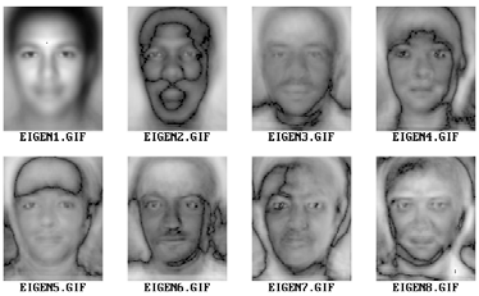

Figure 4. First eight eigenfaces from PCA reported in Tredoux, 1996.

Figure 5. Reconstruction of face images with increasing numbers of eigenfaces (from Tredoux, 2002)

This was a great idea, I thought – as did Peter Hancock, Vicki Bruce, and Mike Burton, working at Stirling University, and a bit later, Stuart Gibson and Chris Solomon in Kent. One interesting application was to use the spatialization of the face images, in the PCA, to derive measures of similarity. Independently, we pursued this application, and came to more or less the same conclusion: there is a moderate or perhaps moderately strong relation between image similarity and similarity of faces rated by human beings. (We have since done some work together on a related application – see Frowd, Bruce, Pitchford, Gannon, Robinson, Tredoux, Park, McIntyre, & Hancock, 2010).

I took this a further step, in the direction I was more interested in, namely the measurement of police lineup fairness. I reasoned that the image similarity measure ought to tell us something about the bias, and effective size of police lineups. In a set of experiments published in 2002, I did just this: I was able to show that there is a significant relationship between the image similarity measure and the bias and effective size of photospread lineups, as assessed with mock witness measures (Tredoux, 2002). In other words, we may be able to assess the fairness and effective size of lineups just by computing image similarity measures on the lineup.

This could also be used for constructing lineups – the technique is not restricted to post hoc assessment of police lineups. The task is as follows: given a suspect, find 5 foils of sufficient similarity to the suspect. Something like this happens now in the UK, where they have switched to an image based lineup technology – they have a large library (over 30 000) of foils, and when police build a lineup they send an image of the suspect to a central office, which uses a computer to search a crude description of the face, and generate a number of candidate foils[10]. The problem with this technique is to have a large enough collection of foils - this is particularly difficult when the suspect is distinctive, or from a population that is not well represented in the database. It would be better to generate artificial or synthetic foils.

This remains a difficult problem, but it may be evident to you from the description of the PCA modeling, that this should be possible. As I finished my PhD in 1995, this struck me as a strong possibility. I started manufacturing faces by randomly generating sets of coefficients, and multiplying them through the eigenfaces, and produced convincing face-like images. I discovered that you could put constraints on the coefficients you generated, but what I did not see was how to control them more systematically.

[10] See the description of the VIPER lineup system at http://www.viper.police.uk/

Good fortune knocked on my door, one day - quite literally! Yon Rosenthal, an MSc student in Electrical Engineering, had been doing just the same thing in an office 200m from me, along with John Greene and Gerhard de Jager. He had also generated synthetic faces, and had found a way of guiding the coefficient search with a hybrid genetic search algorithm (Rosenthal, de Jager, & Greene, 1997; Rosenthal, 1998). He had heard from a friend that I was doing similar things with faces in the Psychology Department, so he walked across the road and knocked on my door.

This was a quite wonderful period of collaboration for me. I had an outstandingly talented research assistant, David Nunez, and the three of us put our heads together. We took Yon's face generation program – then written in Matlab, and based on a very small sample of 32 low resolution faces of British engineering students, available for download from the internet – and worked on just about every aspect of the system Yon had devised. This included collecting over 3000 high resolution images from all over Cape Town.

Most significantly, we recruited the talents of a Grade 12 school pupil, Oliver Oxtoby, who had a reputation even then as being a formidable computer programmer. His first task as a Grade 12 pupil was to read Yon's M.Sc. thesis and to write software that would implement the improvements we had identified! Oliver has since gone on to get a PhD in Mathematics from UCT, but has continued to work with us on an occasional basis. His contributions have been immensely important to our work. Yon and David have since moved on to different parts of the world, Yon early in the project, and David 5 years ago.

I want to fast forward through what has been over a decade of work on these synthetic faces, including a lot of basic work on image capture, and processing. We had to devise our own bespoke algorithms, and we identified better methods for constraining coefficients, based on some statistical theory I was able to dredge up from my own training a long time ago. Along the way, a set of images of average Irish faces that we created with Nuala Brady and Hannah Allen won the Irish Times Science Photography prize (see Figure 6).

Figure 6. 'Average' male and female Irish faces : Winning entry in Irish Times Science Photography competition

{DEMONSTRATION IN THE LECTURE AT THIS POINT}

We use this software to do a number of things, but I want to talk about two only, and these are the two that are germane to the talk tonight. To get back to where this idea came from: can we generate images for use in police lineups that will be acceptable foils? My student Caitlin Grist and I have worked on this in recent years, and we believe that you can.

Figure 7 shows a lineup created to consist of a suspect and five synthetic faces, all of high similarity to the suspect, and Figure 8 shows a lineup created with the same suspect, to be of moderate to low similarity to the suspect.

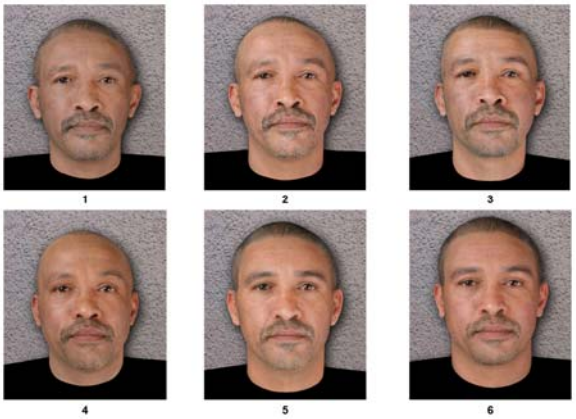

Figure 7. High similarity (suspect – foil) synthetic face lineup.

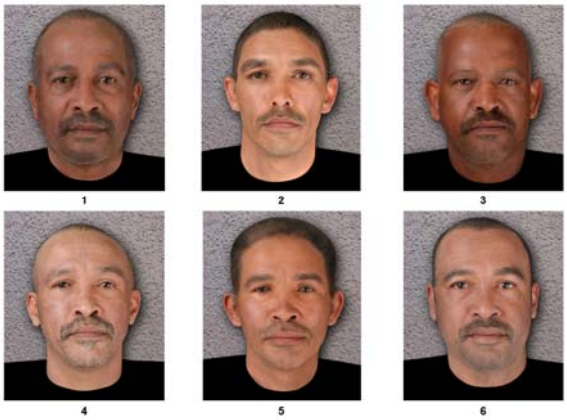

Figure 8. Low similarity (suspect – foil) synthetic face lineup.

What we need to do in this research is to find some tradeoff between the following things: identifications of the suspect when he is guilty, and identifications of the suspect when he is innocent (this is known in the literature as lineup diagnosticity, it is very similar to an index known as relative risk in the biomedical literature). Can we find an optimal level of similarity to do this? Our investigations show that when we manipulate similarity we find predictable increasing levels of this tradeoff – we are not clear at this stage, but we seem to find that more dissimilar foils are better overall, but the unknown cost here is identification of innocent suspects.

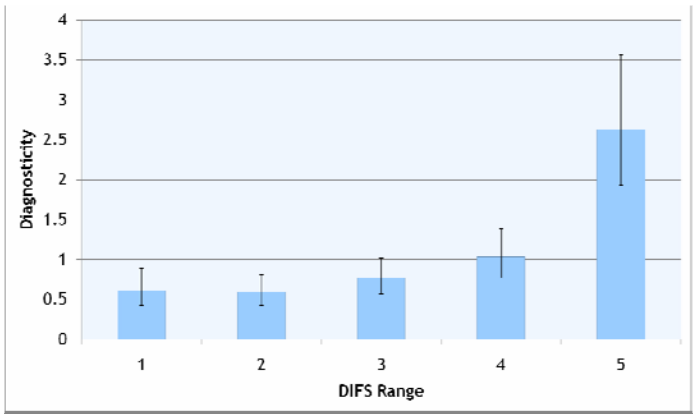

Figure 9. Relationship of lineup suspect-foil dissimilarity (DIFS) to lineup diagnosticity (trade-off of correct identifications of perpetrators to incorrect identifications of innocent suspects)

Changing the medium of the lineup

We have spoken about the way that lineups are constructed, but we have not considered the question of how lineups are conducted. This is something the courts in South Africa have often commented on, and have set down rules about, most of which seem very good.

Social psychologists became interested in the lineup in the late 1970s, particularly in what might are called expectancy effects. The clear problem is that witnesses are too willing to choose someone. Gary Wells (Wells, 1984), and before him, a number of English jurists and law authors (cf. Williams & Hammelmann, 1963), have made this point very simply: if you witness a crime, and after a period of time the police ask you to come to the police station to attempt an identification, you will conclude that the police have a strong case against the suspect, and that your task is to find the person the police suspect. Notice the slant here: you will not look specifically for the person you saw commit the crime, but for the person the police suspect. Wells (1984) dubbed this a 'relative judgement strategy', as opposed to an 'absolute judgement strategy', which is what we really want. Witnesses will set about finding the clues that lead to the solution – is there someone

in the lineup dressed differently? Who has an unusual scar, perhaps? If you follow what the police are doing behind the scenes, you may even be able to see who is led into the room, and from where. The point is that the procedure and situation are highly suggestive, and it will be hard to control these factors perfectly. Our courts recognize this explicitly, and have set down many rules about who should conduct the lineup, and how the witness should be escorted during the procedure, and so on.

A very clear demonstration of how suggestible witnesses are, and how malleable their testimony, is given in a study by Gary Wells, and one of his students, Amy Bradfield (Wells & Bradfield, 1998). In this study, some witnesses were given confirming feedback upon making an identification from a lineup (they were told “Good, you identified the actual suspect"), in contrast to control witnesses who were not given this response. Witnesses given the post-decision feedback were more confident, and showed significant increases in the self-rating of their viewing opportunity, the level of attention they paid, and how quickly they identified the suspect. And obviously, but most importantly, giving witnesses the post-decision feedback did not increase their accuracy relative to control witnesses.

What can we do about the problem that witnesses feel obliged to choose someone from the lineup, and appear to do this by choosing the most suitable candidate among those assembled in the lineup? Gary Wells and Rod Lindsay from Iowa, and Ontario, respectively, have devised all sorts of pragmatic interventions around this.

Wells proposed a ‘blank lineup’ in 1984, arguing that the witness should be tested by preceding the lineup containing the suspect with one in which all the people are innocent. If the witness survives this test, he/she gets to view the lineup which does contain the suspect. This is a good idea, but one that the police have not been particularly interested in, especially in countries where they need to conduct live, or corporeal lineups (since it will involve recruiting foils for two lineups!)

Lindsay & Wells (1985) proposed a more fundamental change to lineup procedure, which is to present the lineup members sequentially, rather than in an array. The idea here is that you cannot compare the foils to each other in any easy manner, and in so doing the relative judgement strategy is countered. They explored this alternative method, and a meta-analysis in 2001 appeared to show that it reduces false identifications of innocent suspects at little, if any cost, to identifications of guilty suspects.

This is a clever idea, and the data supporting it do suggest that there is something better about the sequential lineup. So much so that in several of the United States now, it is obligatory to run sequential rather than simultaneous lineups: New Jersey and North Carolina are two of these states.[11]

But our own research on the sequential lineup shows something different. In a number of studies we show that the sequential advantage is not one of more accurate recognition, but simply a tradeoff (Malpass, Tredoux, & McQuiston-Surrett, 2009; McQuiston-Surrett, Malpass, & Tredoux, 2006; Meissner, Tredoux, Parker, & MacLin, 2005). The number of false positives eliminated with the sequential procedure is directly offset by the number of correct positive identifications that are lost. This may still be preferable – we do not want to convict 1 more guilty person at the cost of falsely convicting 1 innocent person – and is ultimately an issue for public policy, but it does tell us that the sequential procedure is really just a method of making witnesses more conservative. We know from our own data that giving a witness a severe warning at the time of the lineup also does this (Meissner et al., 2005. Our courts have long insisted on a warning to eyewitnesses who are attempting identifications, but it is not clear that the warning is strong enough).

Rod Lindsay and his student, Sean Pryke, have proposed another simple, and very clever practical change to the way in which we conduct lineups (Pryke, Lindsay, Dysart, & Dupuis, 2004). This is known as the multiple channel lineup, and capitalizes on the strong likelihood that you remember more about the perpetrator than just his face. Several lineups are presented to the witness, each consisting of a different view, or feature, of the lineup members. For instance, the first channel might consist of a frontal view of just the faces, the second channel the entire bodies, excluding the faces, the third channel could be voice, and the fourth a video clip of each person walking, but silhouetted so that no details of the body or face are clear. The witness has to make a decision for each lineup. A positive identification of the suspect only occurs when he or she has been selected from all of the lineups. The interesting thing about the multiple channel lineup is what happens when the suspect is innocent, and the witness chooses on the basis of very little information, or just randomly. If we assume that the channels are independent, and if there are six people in the lineup, then a witness who chooses randomly will choose the innocent suspect 1 in 6 times. But since we are only interested in identifications of the suspect from all four lineups, the probability that a randomly choosing witness will do this is $\left(\frac{1}{6}\right)^4$. This event is unlikely to occur. Similar

[11] See http://www.innocenceproject.org/Content/Sequential_Lineups_Are_More_Accurate_According_to_Groundbreaking_Report_on_Eyewitness_Identification_Procedures.php

calculations, assuming the witness' memory is not quite so poor, will show similarly low likelihoods of false positive identifications.

Of course, we also want to know whether this procedure affects identifications from lineups where the perpetrator is present. To re-use the language from earlier in this lecture, we want to achieve good diagnosticity: the ratio of correct identifications to false positive identifications should be high. Empirical tests of this procedure suggest that diagnosticity is good, and that few, if any positive identifications are lost for the large reduction in false positive identifications.

Why does this procedure work? Just as in the case of sequential lineups, we – meaning the research collaborative between my laboratory and researchers in the USA – want to know why a procedure works, not simply that it works. Does it, like the sequential lineup, simply make witnesses more conservative? The results that Pryke et al. report suggests that it does not work by simply making witnesses more conservative, but without an explanation of why, or how, the intervention seems unsatisfactory.

This is the limitation of the pragmatic approach I outlined earlier. I have seen a similar problem in an entirely different literature, namely that concerning the use of the polygraph to detect deception. The key problem with that endeavour is that claims to the validity of the polygraph rest on contested empirical demonstrations that physiological measures can predict deception with reasonable accuracy. The claim is without theoretical substance, though, and without a coherent explanation of why there should be a direct relation between deception and bodily indicators, the entire enterprise flounders (see Tredoux & Pooley, 2001, 2005).

Making faces

Much of what eyewitnesses have to remember is visual, and hard to put into words. In particular, the witness must try to provide information from memory about the perpetrator's face. Police forces have long hankered after recovering visual information about the perpetrator that is better than a verbal description. For many years they employed trained portrait artists to interview witnesses, and to produce interactive sketches. This practice has just about died out, globally: in the first place because it is expensive, and in the second because acetate kits, and computer software, in that order, have taken the place of the sketch artist.

The first of these kits was constructed by Jaques Penry (Penry & Ryan, 1971) from several hundred line drawings of faces. Features were cut out, and stored in catalogues that could be browsed by feature, and the features themselves organized in general categories (e.g. hair colour, length, style).

Later kits were constructed in the same way from large collections of photographs. Psychologists at Aberdeen University – John Shepherd, Gaham Davies, Hadyn Ellis, Rhona Flin, among others – tested the ability of these face composite systems to assist witnesses in producing recognizable portraits, but their results were disappointing. Witnesses were rarely able to produce likenesses that were recognizable at levels greater than mere guessing (Davies, Ellis, & Shepherd, 1978; Davies, 1981; Ellis, Shepherd, & Davies, 1975), even when the faces were in full view at the time of reconstruction.

The advent of inexpensive desktop computers in the 1980s spawned new software face composite programs, and these seemed at first to hold considerable promise. Larger libraries could be assembled, maintained, and searched, for one, but digital editing tools could also be used to disguise the piecemeal nature of the construction, as well as to make substantive modifications to features and to configurations of features. Alas, these programs fared little better, as a review by Graham Davies and colleagues concluded (Davies, van der Willik & Morrison, 2000): what was clear was that software programs allowed better reconstructions for faces directly in view, but not when reconstructed from memory.

Two key problems with these composite systems appeared to be i) their dependency on featural composition, and ii) the inherently limited nature of the facial feature databases they relied on. There is a good deal of literature to suggest that faces are perceived as configurations rather than as collections of features (see Rakover, 2002). It is likely, therefore, that individual features are difficult to match in a contextless task, such as that faced by a witness who has to choose individual features from a library of similar features (e.g. choosing the correct nose for a reconstruction from a collection of noses). The ability of featural compositing systems such as Identikit and Photofit to replace individual features within a facial configuration may ameliorate this problem, but it remains a difficult and apparently unnatural task for witnesses to construct faces from features. Databases of facial features are also inherently limited. Libraries of features will always be limited to the population from which they derive, and the ability of compositions to represent particular faces can only be judged on a trial-and-error basis. The underlying problem with featural face compositors may be the absence of a representational or computational theory. Most compositors are based on the common-sense notion that sectioning a face is invertible - just as a face can be sectioned into components, so it can be recreated by arrangements of sections - but have no theoretical basis which would allow them to answer important questions about their characteristics or abilities. Some of these questions are: With what accuracy does the system represent a randomly chosen member of the population of faces from which the library is drawn? How are the individual features related, for

example in terms of their similarity? What is an unusual and what is a typical face in terms of the set forming the library?

The whole-face generator that we discussed earlier in this talk seems to provide some solutions to these problems. It works only with whole faces, and never with disembodied features. It has the ability to generate an endless number of new faces, but better than all that, it is possible to demonstrate the ability of the face generator to represent members of target populations, and to calculate the amount of error in doing so.

We have put the face generator into use as a face composite system, as have groups in Stirling, Scotland, and Kent, England. Our tests at this stage are in my view inconclusive. It is possible to construct highly realistic and easily recognizable likenesses of faces when they are in direct view, but accuracy from memory – even a well encoded memory – is still poor, but above the level of mere guessing, which is an improvement from where we were in 2000. Figure 10 shows an example of some reconstructions in our laboratory with faces in full view, and Figure 11 reconstructions from memory. These were collected by my student, Taryn Sullivan (Sullivan, 2007).

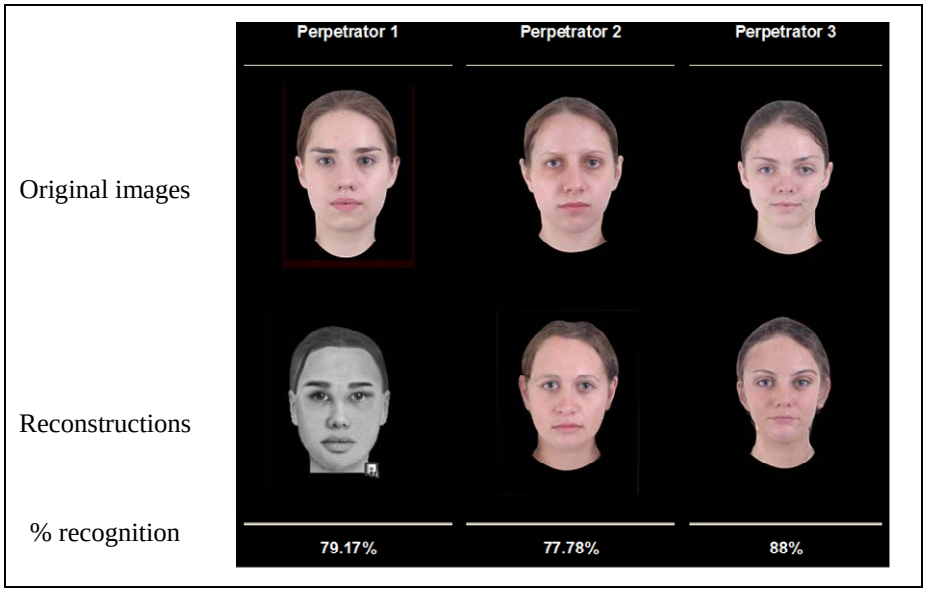

Figure 10. Reconstructions of faces (in view) using FACES (left), and the ID face compositor (centre, and right) (Sullivan, 2007)

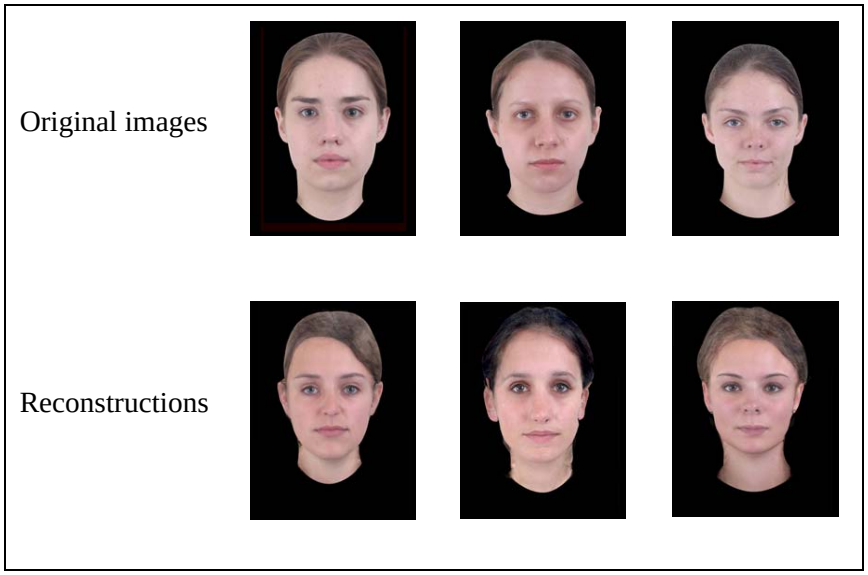

Figure 11. Reconstructions of faces (from memory) using the ID face compositor (Sullivan, 2007)

More radical solutions

The efforts by 'pragmatic psychology' to combat the perils of eyewitness identification have in my view had a moderate level of success. They have certainly roused the interest of the United States Criminal Justice system. Janet Reno, the Attorney General during the Clinton administration, put together a group of eyewitness researchers, lawyers, and police officials, to produce a set of guidelines for criminal justice practice in respect of eyewitnesses: these guidelines cover how their statements should be taken, how they should be interviewed, what identification procedures should be used, and how (see Wells, Malpass, Lindsay, Fisher, Turtle, & Fulero, 2000b; Technical Working Group for Eyewitness Evidence, 1999).

But none of the interventions we have discussed here, or that are published in our journals, reduce the problem to a level where we could ignore it. For instance, the sequential lineup reduces false identifications of innocent suspects to a figure of 9%, against a benchmark of 25% in a control group. This cannot be taken to represent what would happen in sequential lineups conducted by police, in real cases, but it does suggest that even with corrective strategies will not negate the problem, that there will always be a residual proportion of mistaken eyewitness identifications.

Some commentators have seen this, and have pointed to more radical solutions to the problem of eyewitness identification. One solution, which received the strong recommendation of Lord Devlin, in his report of 1976, was to accept eyewitness identifications only when there was corroborating evidence. He argued that experience had shown that it was not safe to convict on the basis of only an eyewitness identification. This was not adopted in England, which is where the report was

published; it is also not clear that it would have helped much in many of the 200 plus US cases where eyewitness evidence was central to the conviction – in about one third of those cases, prosecutors also had sworn confessions from the defendants, which bolstered the cases significantly.

Another approach is to remove it entirely as a form of evidence in criminal trials, but this would depart so fundamentally from how we establish the identity of the defendant as the perpetrator, it seems unlikely to be a solution.

Conclusion, and the case of Afzal Simons

Rather than concluding with a concise version of what I have already said, I want to end by taking up a particular South African case. This is the case of Afzal Norman Simons[12], who is presently serving a term of 40 years, or life, in prison for the murder of an 11 year old boy, Elroy van Rooyen, in March 1994. The young boy was murdered in a manner that was very similar to 21 other murders in the Western Cape, in the period 1986 – 1994, and which were ascribed in the press and by the police to a common perpetrator dubbed 'the Station Strangler'.

The case has many layers to it. Firstly, it is important to remember the context. Someone had been raping and ritually murdering young boys with impunity, for over 8 years, binding their hands behind their backs, and pushing their heads into the sands of the Cape Flats. The police had made no headway on this case despite forensic evidence available from the many corpses the murderer left behind, and despite multiple eyewitness sightings, and face composite sketches. The murders picked up pace in the early months of 1994, with the corpses of 11 boys discovered in January, alone. Usually the corpses were discovered by members of the public, by accident/fortuitously. People on the Cape Flats and indeed in the whole of Cape Town were outraged: in one incident, a crowd of people stormed the Steenberg police station when they heard the rumour that the police had arrested the Strangler.

At the same time, one should remember that Apartheid was being dismantled, and the first democratic elections of April 27, 1994 were just a month away. That a very large police operation was mounted, under the control of Colonel Leonard Knipe, and a substantial reward of R250 000 offered for information leading to the arrest of the Strangler, was seen by many as at least in part a

[12] Details of the case were recovered from Pistorius, 2000; Pistorius, 2002, Harber, 2010, several newspaper articles published by Lavern de Vries and Warda Meyer, a repository of articles available at http://www.smartcape.org.za/articles/index.php?dir=General/Station%20Strangler/Station%20Strangler%20, as well as the tranche of documents available in the Simons case file, Archives of the High Court, Cape Town.

gesture aimed at gaining political support in the Western Cape among the residents of the Cape Flats.

Simons was a patient at a local psychiatric clinic in April 1994, where a nurse noticed his resemblance to a face composite sketch the police had released. She notified the police, and they followed Simons from the clinic, and asked him to accompany them to their mobile police unit. He spent the night there, after what he assumed was a casual conversation, and the police arrested him the next morning when he woke.

He was placed on an identification parade two days later, and two witnesses were brought to it to attempt an identification. One witness was a young boy – in fact, the cousin of the victim - who had accompanied the victim and perpetrator, but who became suspicious and had run to safety. The other was a woman, Mrs Fouzia Hercules, who had seen the perpetrator walk away from a shopping centre in the company of two young boys. The young boy was unable to identify anyone from behind the one way mirror, but Mrs Hercules identified Simons after a fairly lengthy and hesitant inspection of the lineup (see Figure 12).

Figure 12. Screenshot from lineup in the Afzal Simons case. Simons is number 23.

This identification turned out to be crucial at the trial, as did the purported similarity between Simons and the composite sketch. Figure 13 shows a photograph of Simons taken on the day van Rooyen was murdered, and several composite sketches produced during the case by witnesses.

Figure 13. Face composite sketches made by eyewitnesses in the 'Station Strangler' case

The identification took place on the Thursday, and Simons was taken to a magistrate on the Saturday to make a confession. To what extent this really was a confession, and to what extent Simons' mental state coloured the confession with delusions, is contentious. He later stated in an affidavit that he had made the confession under duress, and in a confused state of mind.

On the following Monday early in the morning (pre-dawn) Simons was taken by Police Lieutenant Barkhuizen to point out sites where he had allegedly left corpses of boys he had murdered. It appears from the court records we have managed to get hold of that he made one pointing out, of unclear relevance, before declaring that he could make no more.

But particularly alarming about this case is that a fair amount of biological evidence was recovered from several of the Strangler victims, but not from Elroy van Rooyen. This biological evidence points to someone else, not to Simons. In particular, semen samples recovered from two of the murder scenes matched each other, but did not match Simons.

Simons protested his innocence at his trial, where he was sentenced to 25 years in prison. He was represented at that trial by Adv Koos Louw. His case went on appeal to Bloemfontein, where he was represented by Adv Marius Broeksma. His appeal was rejected, and his sentence increased to 40 years. Both advocates are adamant that Simons is innocent, and they insist this beyond what might be construed as their professional duty to defend him.

What is there to say in the case of Afzal Norman Simons? There are striking similarities to many of the cases in the United States where DNA evidence has exonerated 273 people. The first is eyewitness evidence of questionable merit. One eyewitness was unable to make an identification, and another made a hesitant identification ("Ek dink dis hy"), both from a lineup that seems to have made insufficient effort to ensure that all members were dressed similarly. A superficial resemblance to a face composite was taken into account by a trial judge. A confession was submitted as evidence, but there are serious questions about the validity of the confession, and whether it is a confession at all. There are also lingering concerns about the mental status of Simons during the interrogation. Most concerning is the fact that although police at the time assumed the murders were the acts of one person, biological evidence at some of the murder sites does not match Simons.

==========

A final comment, rather than a summary. I think this has been a fascinating period of meddling in public life for eyewitness and face recognition researchers, but I think the limits of this gumboots and all approach are clear to me. We do need theoretical accounts of what goes on, at least in the proposed interventions that we make. In other words, a theory not of memory and face recognition, not a universal theory of how the mind apprehends and recollects visual phenomena, but a small scale theory of why changes in the procedures we use to elicit identifications should yield more accurate results.

References

Bartlett, F. C. (1932). *Remembering*. Cambridge, England: Cambridge University Press.

Binet, A. (1900). La suggestibilité [on suggestibility]. *Paris: Schleicher.*

Borchard, E. (1932). *Convicting the innocent: Errors of criminal justice.* New York: Da Capo Press.

Brennan, S. E. (1982). *The caricature generator.* (Unpublished MSc.). Massachusetts Institute of Technology.

Brennan, S. E. (1985). Caricature generator: The dynamic exaggeration of faces by computer. *Leonardo, 18*(3), 170-178.

Coates, T. (1999). *The strange story of Adolph Beck.* London: Stationery Office Press.

Danziger, K. (1990). The autonomy of applied psychology. *22nd International Congress of Applied Psychology.* Kyoto, Japan.

Davies, G. M. (1981). Face recall systems. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and Remembering Faces* (pp. 227-250). London: Academic Press.

Davies, G., Ellis, H. D., & Shepherd, J. (1978). Face identification - the influence of delay upon accuracy of photofit. *Journal of Police Science and Administration, 6*, 34-42.

Davies, G.M., van der Willik, P., & Morrison, L.J. (2000). Facial composite production: A comparison of mechanical and computer-driven systems. *Journal of Applied Psychology, 85*, 119-124.

Devlin, P. D. (1976). *Report to the secretary of state for the home department of the departmental committee on evidence of identification in criminal cases.* London: Her Majesty's Stationery Office.

Doob, A. N., & Kirshenbaum, H. M. (1973). Bias in police lineups — partial remembering. *Journal of Police Science and Administration, 18*, 287-93.

Ellis, H., Shepherd, J., & Davies, G. (1975). An investigation of the use of the photo-fit technique for recalling faces. *British Journal of Psychology, 66*(1), 29-37.

Fisher, R. P., McCauley, M. R., & Geiselman, R. E. (1994). Improving eyewitness testimony with the cognitive interview. In D. Ross, J. Read, and M. Toglia (eds.), *Adult Eyewitness Testimony: Current Trends and Developments*. (pp. 245-69). London: Cambridge University Press

Fisher, R. P. (1995). Interviewing victims and witnesses of crime. *Psychology, Public Policy, and Law, 1*(4), 732-764.

Fisher, R. P., & Schreiber, N. (2007). Interview protocols for improving eyewitness memory. In R. C. L. Lindsay (Ed.), *The Handbook of eyewitness psychology, vol I: Memory for events.* (pp. 53-80). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Frowd, C.D., Bruce, V., Pitchford, M., Gannon, C., Robinson, M., Tredoux, C., Park., J., McIntyre, A., & Hancock, P.J.B. (2010). Evolving the memory of a criminal's face: methods to search a face space more effectively. *Soft Computing, 14*, 81-90.

Geiselman, R. E., Fisher, R. P., Mackinnon, D. P., & Holland, H. L. (1986). Enhancement of eyewitness memory with the cognitive interview. *American Journal of Psychology, 99*(3), 385-401.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66*(3), 325-331.

Gombrich, E. H., Hochberg, J. E., & Black, M. (1973). *Art, perception and reality.* Baltimore: Johns Hopkins University Press.

Harber, A. (2010). *Troublemakers: The best of South Africa's investigative journalism* Jacana Media.

Ihde, D. (2010). *Heidegger's technologies: Postphenomenological perspectives.* New York: Fordham University Press.

Jeffreys, A. J., Wilson, V., & Thein, S. L. (1985). Individual-specific `fingerprints' of human DNA. *Nature, 316*(6023), 76-79.

Junkin, T. (2004). *Bloodsworth: The true story of the first death row inmate exonerated by DNA.* Algonquin Books of Chapel Hill.

Kirby, M., & Sirovich, L. (1990). Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 12*, 103-108.

Lindsay, R. C. L., Mansour, J. K., Beaudry, J. L., Leach, A., & Bertrand, M. I. (2009). Beyond sequential presentation: Misconceptions and misrepresentations of sequential lineups. *Legal and Criminological Psychology, 14*(1), 31-34.

Lindsay, R. C. L., & Wells, G. L. (1985). Improving eyewitness identifications from lineups - simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70*(3), 556-564.

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

Malpass, R.S., & Devine, P.G. (1981). Guided memory in eyewitness identification. *Journal of Applied Psychology, 66*, 343-350.

Malpass, R. S., Tredoux, C. G., & McQuiston-Surrett, D. (2009). Public policy and sequential lineups. *Legal and Criminological Psychology, 14*(1), 1-12.

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*(2), 137-169.

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition, 33*(5), 783-792.

Mullis, K. B., & Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. In Ray Wu (Ed.), *Methods in enzymology* (pp. 335-350). New York: Academic Press.

Munsterberg, H. (1908). *On the witness stand: Essays on psychology and crime*. New York: Doubleday

Penry, J., & Ryan, I. (1971). *Looking at faces and remembering them: A guide to facial identification.* London: Elek Books.

Pistorius, M. (2000). *Catch me a killer*. Penguin Global.

Pistorius, M. (2002). *Strangers on the street: Serial homicide in South Africa*. Penguin Global.

Potter, J. (1982). "... Nothing so practical as a good theory." The problematic application of social psychology. In P. Stringer (ed), *Confronting social issues***.** (pp 23-49). London: Academic Press.

Pryke, S., Lindsay, R. C. L., Dysart, J. E., & Dupuis, P. (2004). Multiple independent identification decisions: A method of

calibrating eyewitness identifications. *Journal of Applied Psychology, 89*(1), 73-84.

Rakover, S. S. (2002). Featural vs. configurational information in faces: A conceptual and empirical analysis. *British Journal of Psychology, 93*, 1-30.

Rembar, C. (1980). *The law of the land: The evolution of our legal system.* Simon and Schuster.

Rosenthal, Y., de Jager, G., & Greene, J. A computerized face recall system using eigenfaces. Proceedings of the Eighth Annual South African Workshop on Pattern Recognition (pp. 53–57). Rhodes University: Grahamstown.

Rosenthal, Y. (1998). A computerised face recall system using eigenfaces. Unpublished M.Sc. thesis, University of Cape Town.

Rust, A., & Tredoux, C. G. (1998). Identification parades: An empirical survey of legal recommendations and police practice in south africa. *South African Journal of Criminal Justice, 11*, 196-212.

Scheck, B., Neufeld, P., & Dwyer, J. (2001). *Actual innocence.* New York: Doubleday.

Shepherd, J., Ellis, H. D., & Davies, G. (1982). *Identification evidence: A psychological evaluation.* Aberdeen: Abderdeen University Press.

Sirovich, L., & Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America, 4*(3), 519-524.

Sporer, S. L. (2008). Lessons from the origins of eyewitness testimony research in europe. *Applied Cognitive Psychology, 22*(6), 737-757.

Sporer, S. (1982). A brief history of the psychology of testimony. *Current Psychological Reviews, 2*, 323–339.

Stern, W. (1910). Abstracts of lectures on the psychology of testimony and on the study of individuality. *The American Journal of Psychology, 21*(2), 270-282.

Sullivan, T. (2007). Comparing featural and holistic composite systems with the aid of guided memory techniques. Unpublished M.A. thesis, University of Cape Town.

Technical Working Group for Eyewitness Evidence (1999). *Eyewitness evidence: A guide for law enforcement* (NCJ No. 178240). Washington, DC: National Institute of Justice.

Thompson-Cannino, J., Cotton, R., & Torneo, E. (2009). *Picking Cotton: Our memoir of injustice and redemption.* New York: St Martins Press.

Tredoux, C.G. & Pooley, S. (2001). Polygraph testing of deception: an evaluation and commentary. *Industrial Law Journal, 22*, 819-839

Tredoux, C. G., & Tredoux, P. J. (1987, September). Identification parades in South African law: A psychological evaluation. *7th Annual Congress of the Psychological Association of South Africa*, Cape Town, South Africa.

Tredoux, C. G. (1996). *Evaluating the fairness of identification parades with measures of facial similarity.* Unpublished Ph.D. thesis, University of Cape Town.

Tredoux, C.G. (1998). Applied Psychology: application of psychological knowledge or nominalist error? In J. Mouton & J. Muller. (Eds) *Knowledge, Method and the Public Good*. Pretoria: HSRC Press

Tredoux, C. G. (1998b). Statistical inference on measures of lineup fairness. *Law and Human Behavior, 22*(2), 217-237.

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

Tredoux, C. G. (2002). A direct measure of facial similarity and its relation to human similarity perception. *Journal of Experimental Psychology: Applied, 8*(3), 180-193.

Tredoux, C. G., & Chiroro, P. M. (2005). Eyewitness testimony. In C. G. Tredoux, D. H. Foster, A. Allen, A. Cohen & D. Wassenaar (Eds.), *Psychology and law* (pp. 193-225). Cape Town: Juta.

Tredoux, C. G., & Pooley, S. (2005). Credibility assessment. In C. G Tredoux, D. H. Foster, A. Allan, A. Cohen, & D. Wassenaar. (Eds). *Psychology and law*. Cape Town: JUTA.

Tredoux, C., Parker, J. F., & Nunez, D. (2007). Predicting eyewitness identification accuracy with mock witness measures of lineup fairness: Quality of encoding interacts with lineup format. *South African Journal of Psychology, 37*(2), 207-222.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352.

Valentine, T. (1991). A unified account of the effects of distinctiveness inversion and race in face recognition. *The Quarterly Journal of Experimental Psychology, 43a*(2), 161-204.

Verwoerd, H. F. (1929). A contribution to the experimental investigation of testimony. *South African Journal of Science, 26*, 949-957.

Wells, G. L. (1978). Applied eyewitness-testimony research - system variables and estimator variables. *Journal of Personality and Social Psychology, 36*(12), 1546-1557.

Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology, 14*(2), 89-103.

Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station - A successful application of eyewitness research. *American Psychologist, 55*(6), 581-598.

Wells, G. L., & Bradfield, A. L. (1998). 'Good, you identified the suspect': Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology, 83*(3), 360-376.

Williams, G., & Hammelmann, H. A. (1963). Identification parades. *Criminal Law Review*, 479-91,545-555.

Wright, C. A., & Graham, K. W. (2005). Federal practice & procedure: Evidence. St Paul: Minnesota.

Zeffertt, D., Paizes, A., & Skeen, A. S. Q. (2003). *The South African law of evidence*. Durban: Butterworths.