

ESSAYS IN CRIMINAL
AND FORENSIC
PSYCHOLOGY NO. 6

KEVIN BREWER

ISBN: 978-1-904542-61-2

This document is produced under two principles:

1. All work is sourced to the original authors. The images are all available in the public domain (most from http://commons.wikimedia.org/wiki/Main_Page). You are free to use this document, but, please, quote the source (Kevin Brewer 2010) and do not claim it as you own work.

This work is licensed under the Creative Commons Attribution (by) 3.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/> or send a letter to Creative Commons, 171 2nd Street, Suite 300, San Francisco, California, 94105, USA.

2. Details of the author are included so that the level of expertise of the writer can be assessed. This compares to documents which are not named and it is not possible to tell if the writer has any knowledge about their subject.

Kevin Brewer BSocSc, MSc

An independent academic psychologist, based in England, who has written extensively on different areas of psychology with an emphasis on the critical stance towards traditional ideas.

(<http://kmbpsychology.jottit.com>)

Orsett Psychological Services,
PO Box 179,
Grays,
Essex
RM16 3EW
UK

orsettpsychologicalservices@phonecoop.coop

CONTENTS

	Page Number
1. False Confessions are Easy to Obtain?	4
2. Accurately Identifying the Faces of Criminals	19
3. Measuring Malingering in the Criminal Justice System	31
4. Genes and Crime, and the Complexity of the Relationship	38
5. Femicide as a Public Health Issue	42

1. FALSE CONFESSIONS ARE EASY TO OBTAIN?

- 1.1. Introduction
- 1.2. Police officers' behaviour
- 1.3. Young age
 - 1.3.1. Evaluation of Gudjonsson et al (2009a)
- 1.4. False conviction
- 1.5. Appendix 1A - Kassin and Kiechel (1996)
- 1.6. References

1.1. INTRODUCTION

Police interrogation aims to gain information from suspects, including ideally a confession of guilt. But just because a suspect confesses to the offence does not mean they actually did it. There are cases of false confessions.

Yet there is a commonsense belief that individuals do not confess to crimes they did not commit, unless they have clear ulterior motives like protecting another person. However, DNA evidence shows that up to a quarter of prisoners who confessed to the police were not guilty (Kassin 2008) ¹.

Kassin and Wrightsman (1985) suggested three different types of false confession:

i) "Voluntary" - Individuals voluntarily go to the police to confess, and there is no police pressure. The reasons include a need for attention; strong general feelings of guilt and/or a need for self-punishment; the individuals are unable to distinguish fact from fantasy (delusions); the perception of tangible gain (eg: fame/infamy); or confessing to protect the real offender.

This is especially so in high-profile cases. For example, over fifty people confessed to the murder of Elizabeth Short in 1947 (known as the "Black Dahlia" murder) (Kassin 2008).

ii) "Coerced-compliant" - Here individuals confess because of the pressure during the interrogation. The pressure may involve the belief of being allowed to go home after confessing, or meeting the implicit demands of the interview. The suspect still knows they did not commit the crime.

For example, five teenagers confessed to the rape and assault of the "Central Park jogger" (a 28-year-old woman) in New York City in 1989 after lengthy

¹ Self-reported false confessions by prison inmates may be around 10% (Horselenberg et al 2003).

interrogation, based on the belief that they would go home then. The conviction was overturned by DNA evidence in 2002 showing Matias Reyes (a convicted rapist and murderer) guilty (Russano et al 2005).

iii) "Coerced-internalised" - Suspects come to believe during the interview that they did commit the offence. This phenomena is based on "memory distrust syndrome", where the individual starts to doubt their own memory of events (Gudjonsson and MacKeith 1982)². Alternatively the suspect may have amnesia at the beginning of the interview.

For example, teenager Michael Crowe came to believe, after lengthy interrogation, that he had murdered his sister. He said, "I'm not sure how I did it. All I know is I did it" (Drizin and Colgan 2004). "Eventually, he was convinced that he had a split personality - that 'bad Michael' acted out of jealous rage while 'good Michael' blocked that incident from consciousness" (Kassin 2008 p249). The real murderer was later found with the sister's blood on his clothing.

Gudjonsson (2003) proposed a series of factors that occur in the interrogation process to explain false confessions:

- Contextual factors - eg: seriousness of crime, strength of police evidence.
- Custodial factors - related to length and nature of the interrogation.
- Vulnerability factors - eg: mental illness, learning disability, suggestibility, young age.
- Support factors - eg: presence of a legal advisor during the interrogation.

Furthermore, Kassin (2008) observed that:

Recent DNA exonerations reveal three sets of problems with confession evidence: (a) Police cannot accurately distinguish between truth tellers and liars; (b) certain psychological interrogation tactics put innocents at risk to confess, especially if they are young, mentally impaired, or otherwise vulnerable; and (c) judges and juries intuitively tend to trust confessions, even if they know that these confessions were coerced (p252).

² This is also similar to "false memory syndrome" (Horselenberg et al 2003).

1.2. POLICE OFFICERS' BEHAVIOUR

There are a number of ways that police officers can encourage false confessions by their behaviour during interrogation.

For example, police officers may be over-confident about the guilt of the suspect and/or their ability to detect deception. Detection of deception is about 54% accurate in experiments, even with training. In fact, police officers in the USA who underwent training to detect deception were less accurate, but more confident of their ability, than those not trained (Kassin 2008).

In the USA the police are permitted to tell suspects that they have evidence of their guilt (eg: fingerprints, witnesses) when none exists (Kassin 2008). Such misinformation can even alter people's memories during interrogation, as shown in an experiment by Kassin and Kiechel (1996) (appendix 1A).

Participants were given a reaction time test, but told not to touch the "ALT" key on the computer keyboard. In some cases, there was a witness who made the false claim of seeing the participants touch the "ALT" key. On average 69% of the participants admitted touching the "ALT" key after this false claim, with 30% admitting so to another person, and 10% could picture themselves doing so.

The authors concluded "that the presentation of false incriminating evidence - an interrogation ploy that is common among the police and sanctioned by many courts - can induce people to internalise blame for outcomes they did not produce" (Kassin and Kiechel 1996 p127).

Another tactic used in police interrogation is known as minimisation³. The interrogator offers sympathy and moral justification to the suspect suggesting that the actions were accidental or provoked, for example. Such behaviour can lead the suspect to believe that a confession will produce leniency.

Russano et al (2005) showed this effect in a laboratory experiment. Individual participants were paired with a confederate to solve problems, but told to work alone and not communicate. In the "innocent" condition, these instructions were followed. In the "guilty" condition, the confederate asked for help. After both conditions the experimenter accuses the participant

³ The self-reported frequency of minimising the moral seriousness of the offence by 631 North American detectives was 3.02 out of 5 (where 5 = always use and 1 = never use) (Kassin et al 2007). This was tenth most frequent with "isolating the suspect from family and friends" and "conducting interrogation in a small private room" being the most popular. Ninth most frequent was "pretending to have independent evidence of guilt".

of cheating and asks them to sign a confession. The students received course credit for participation and this was used to encourage the confession.

Four different tactics were used to achieve this - explicit promises of leniency (deal) ("things could probably be settled pretty quickly"), minimisation ("I'm sure you didn't realise what a big deal it was"), combination of both, or none. This produced eight conditions (figure 1.1).

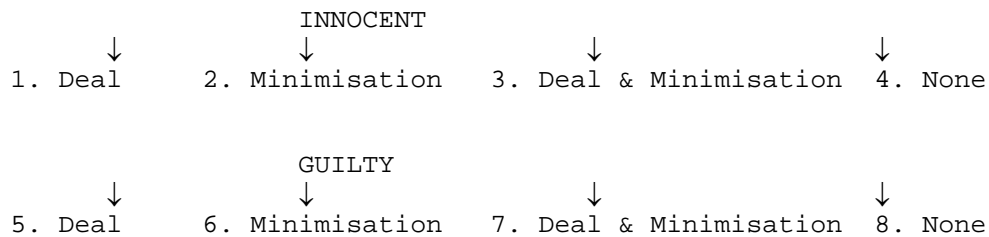


Figure 1.1 - Conditions in experiment by Russano et al (2005).

Willingness to confess was highest in the combination condition (over 80% of true confessions and 40% of false confessions), followed by the minimisation condition (80% and 20% respectively) ahead of explicit leniency. Minimisation, thus, increased confessions by the guilty and innocent (true and false confessions) (figure 1.2).

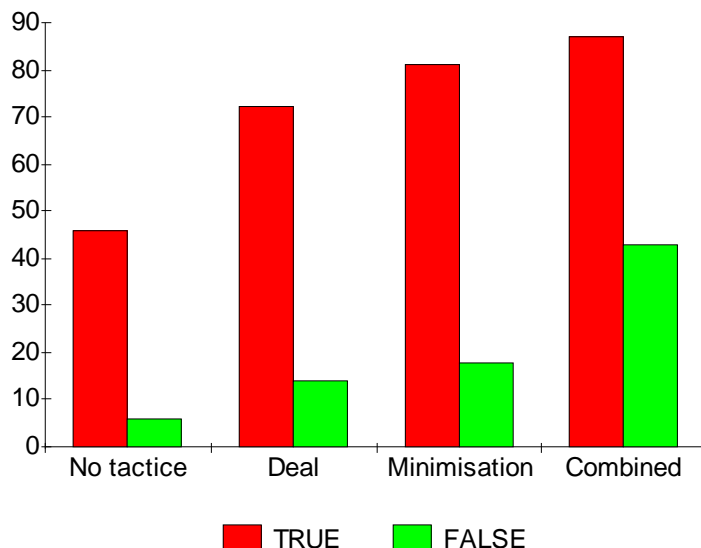


Figure 1.2 - Rates of confession (%) based on tactics used in interrogation.

1.3. YOUNG AGE

Young age is a key vulnerability factor in false confessions. Many adolescent suspects show "immaturity of judgment" which can produce a disproportionate number of false confessions. This includes impulsivity, and focus on immediate gratification without thoughts of the consequences of actions (Kassin 2008). So a strong possibility of confessing as a means to get out of the interrogation situation and go home.

A large scale survey of high school students in various European countries has found self-reported false confessions by 12.2% of those questioned by the police (Gudjonsson et al 2009b) ⁴. In such studies, victims of bullying, for example, are much more likely to falsely confess (Gudjonsson et al 2008).

Gudjonsson et al (2009a) investigated general victimisation of young people (in relation to violence, sexual abuse, serious illness or accident, death of a close relative, parental separation, and history of rejection) and false confession to the police among 7149 15-16 year-olds in Iceland.

They were asked how often they had been interrogated by the police, at a police station, on suspicion of criminal involvement, and whether they had admitted during interrogation to an offence that they had not committed. A five-point scale was used to score the answers. The remainder of the questionnaire, over one hundred questions, asked about topics like family background, personality, and adverse life events including fourteen victimisation experiences (which were divided into four categories) (table 1.1).

VIOLENCE

- Had serious argument with parents
- Witnessed parents' serious argument
- Witnessed serious violence at home involving adults
- Experienced serious violence at home involving adults
- Experienced physical violence

SEXUAL ABUSE

- Sexually abused by an adult within family
- Sexually abused by an adult outside family

SERIOUS ILLNESS, ACCIDENT OR DEATH OF A FAMILY MEMBER OR FRIEND

- Parent or sibling died
- Friend died
- Experienced serious accident

⁴ Seven countries and nine cities - Norway (Oslo), Finland (Helsinki), Iceland (Reykjavik), Latvia (Riga), Lithuania (Vilnius, Kaunas, Klaipedia), Russia (St. Petersburg), and Bulgaria (Sofia).

- Experienced serious illness

SEPARATION/REJECTION

- Parents separated or divorced
- Rejected by friends
- Expelled from school

Table 1.1 - Victimization experiences used in Gudjonsson et al (2009a).

Overall, 758 pupils admitted to experience of police interrogation (11% of sample), and 84 of those (11.3%) claimed to have made a false confession (1.2% of total sample).

Of the fourteen variables of victimisation, only two ("had serious argument with parents" and "experienced serious accident") were not significantly related to false confessions. The strongest relationships with individual variables were "sexually abused by an adult within family" (over six times more likely to be false confessor than not), and "parent or sibling died" (four times more likely) (figure 1.3).

When the variables were combined, "witnessed parents' serious argument", "experienced physical violence", and "parent or sibling died" together explained the largest difference between false and non-false confessors.

MORE LIKELY - FALSE CONFESSION

↑

6.73	sexually abused by an adult within family *
4.13	parent or sibling died *
3.61	sexually abused by an adult outside family *
3.36	witnessed serious violence at home involving adults *
3.15	experienced serious violence at home involving adults *
2.80	witnessed parents' serious argument *
2.58	expelled from school *
2.46	friend died *
2.39	experienced physical violence *
2.14	parents separated or divorced *
2.00	rejected by friends **
1.81	experienced serious illness ***
1.30	had serious argument with parents ****
1.01	experienced serious accidents ****

↓

MORE LIKELY - NOT

(* p<0.001, ** p<0.01, *** p<0.05, **** not significant)

Figure 1.3 - Odds ratios of being false confessor with victimisation variables.

1.3.1. Evaluation of Gudjonsson et al (2009a)

1. The study was based on self-reports of experiencing police interrogation and of giving a false confession with no independent verification of information. Individuals vulnerable to false confession may have admitted to it on the questionnaire when they had not in real life. This could make the figures overestimates.

On the other hand, pupils may have not admitted to interrogation or confession, and so the figures are underestimates.

However, the rate of false confession is similar to other studies which suggests there was not an underestimation or overestimation.

Furthermore, independent verification would be difficult because anonymity would be lost, and the participants may have been less willing to admit to false confessions in that situation. Overall 252 pupils did not answer the question about experiencing police interrogation.

2. The study was based on self-reports of life experiences and victimisation variables. Any self-reported questionnaire depends on honesty as well as accurate recall of information.

3. No details were collected of the nature of the offences falsely confessed to, nor of the reasons for it.

4. The questionnaires were completed in scheduled classes by the volunteers, and took about one hour to complete. There were over one hundred questions and the participants may have become tired or bored with so many to answer.

There is also a question of who did not volunteer to participate. Could they be more or less likely to have experienced police interrogation than the volunteers?

5. Many of the questions were scored on frequency scales (eg: never, once, twice, three-five times, and six or more times), which were reduced to binary variables for analysis. The decision of where to put the cut-off point can influence the findings. Should the cut-off point be never, once, and twice as one variable, and 3-5 times and six or more times as the other, for example, or never and once versus the other options?

6. The use of certain terms are open to interpretation and thus a different response - eg: "serious". What is serious to one person may not be so to another and so on.

1.4. FALSE CONVICTION

Innocent individuals may show behaviour that ironically leads to their false conviction: "People who stand falsely accused believe that truth and justice will prevail and that their innocence is transparent to others. As a result, they cooperate with police, waive their rights, and speak freely, often not realizing that they are under suspicion" (Kassin 2008 p251).

Kassin and Norwick (2004) set up an experiment involving a "mock theft" of \$100. Half of the participants were instructed to take the money left in a room (guilty) and the other half not (innocent). Then the participants were "arrested" and questioned. Of the innocent participants, 81% signed a waiver of their "Miranda rights"⁵ and talked freely in the interrogation when told they did not have to say anything, and this compared to 36% of guilty participants.

Juries pay attention to confessions, even if they are shown to be elicited under pressure. In a mock jury experiment, Kassin and Sukel (1997) found a much higher conviction rate for confessions (even when clearly produced under duress) than with no confessions (47% vs 19%).

1.5. APPENDIX 1A - KASSIN AND KIECHEL (1996)

Seventy-nine undergraduates at a college in the USA took part in what appeared to be a reaction time experiment. They were asked to type letters on a computer keyboard as quickly as the letters were read out by the experimenter. But before this began, the participants were told not to press the "ALT" key as this would cause the computer to crash. After sixty seconds of the experiment the computer automatically crashed, and the angry experimenter asked, "Did you hit the 'ALT' key"?. Would the participant admit to the "crime" which they had not done?

There were two independent variables (IV) and four independent conditions (table 1.2). One IV was the participant's level of vulnerability, which was manipulated by varying the speed of reading the letters.

⁵ These, in the USA, are the rights to silence and to a lawyer.

In the slow-pace condition, 43 letters were read out per minute compared to 67 per minute in the fast-pace condition.

The second IV related to false incriminating evidence. As the experimenter read out the letters to the individual participant, there was a female confederate nearby (a woman who appeared to be a participant waiting her turn, but was working for the experimenter). In the false-witness condition, the confederate said that she saw the participant hit the "ALT" key. In the no-witness condition she said nothing.

Fast-pace/No-witness	Fast-pace/False-witness
Slow-pace/No-witness	Slow-pace/False-witness

Table 1.2 - The four conditions in the experiment by Kassin and Kiechel (1996).

The dependent variable was admitting to the "crime", and was measured in three ways:

i) Compliance - The participant wrote and signed a standard confession, "I hit the 'ALT' key and caused the program to crash. Data were lost". The consequence of this action was a telephone call from the principal investigator. In this case, the participant may admit to the "crime" just to please the experimenter, and not believe that they had pressed the 'ALT' key. This is a coerced-compliant false confession.

ii) Internalisation - After the signed confession scenario (whether the participant had confessed or not), another confederate asked the participant what had happened to see if the participant would confess to the "crime". The conversation was recorded, and then analysed for any wording of a confession (eg: "I hit the wrong button and ruined the program"). Confessions here show that the participant had come to believe that they were guilty (coerced-internalised false confession).

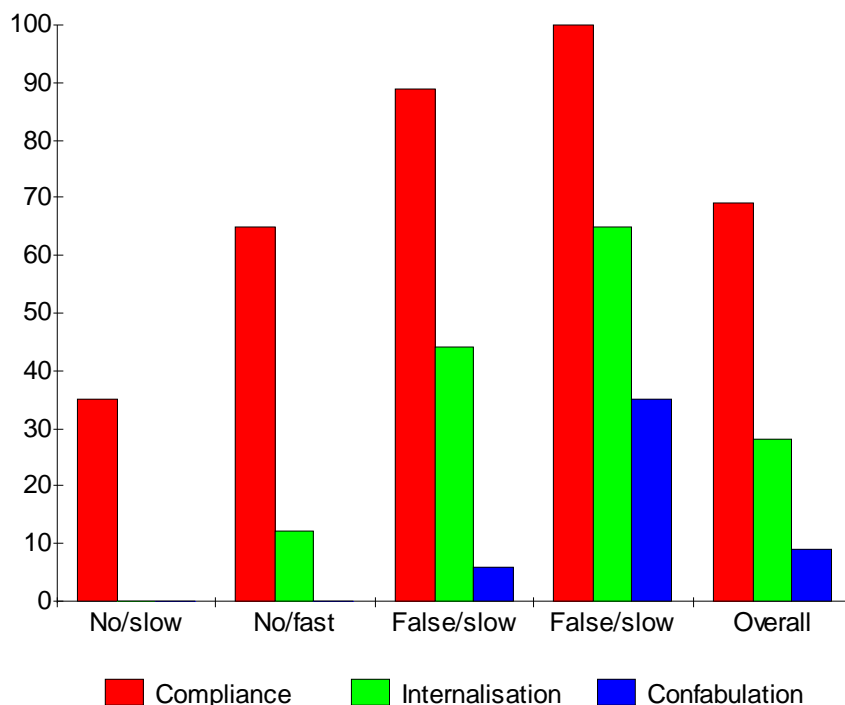
iii) Confabulation - The experimenter asked the participant to re-enact the experiment and show when they hit the "ALT" key (eg: "I hit it when you called out the 'A'"). This tested whether the participant's actual recall was altered by the false witness. This is another aspect of coerced-internalised false confessions.

In the post-experiment debriefing, all but four participants said they believed the scenario as they were told the truth. The authors admitted that most participants showed relief (that they had not ruined the

experiment), amazement (at being manipulated) or satisfaction (at taking part in the research).

Overall, 69% of participants complied (ie: signed confession), but this was all participants in the fast-paced false-witness condition. This was a significant difference compared to the no-witness conditions. Of the participants in the no-witness conditions who signed the confession, none showed internalisation or confabulation suggesting coerced-compliant confessions.

There were seventeen participants in the fast-pace false-witness condition, with 65% showing internalisation and 35% confabulation. The false witness produced significantly more false confessions than no witness (figure 1.4).



(No = no-witness; false = false-witness; slow = slow-pace; fast = fast-pace)

Figure 1.4 - Percentage of participants who showed three types of confession.

Evaluation

There are two main issues with this laboratory experiment by Kassin and Kiechel - validity, and ethical issues.

1. Validity

This relates to whether the experiment and its

findings are applicable to real-life situations. The authors felt that the principle of coerced-internalised false confessions was established for individuals in vulnerable situations faced with false incriminating evidence, though the situation was artificial and the consequences of confession small. The task was artificial, but the situation of being falsely accused was applicable to real-life.

Russano et al (2005) disagreed because all the participants were innocent of what accused (compared to some guilty in real-life), and the participants were only accused of accidentally committing the "crime" (whereas the police see suspects as deliberately committing the offence).

Horselenberg et al (2003) argued that the experimental scenario was not the same as real-life situations for two reasons:

- The confederate giving the false witness claimed to be independent, but in real-life false incriminating evidence would be provided by vested interests like the police.
- Confession had little negative consequences.

To overcome these two weaknesses, Horselenberg et al (2003) replicated Kassin and Kiechel's experiment with the following changes:

a) The experimenter was the false witness (ie: said they had seen the "crime" with their own eyes).

b) A financial loss for confessing (ie: only \$2 fee given instead of \$10 for participating in the experiment).

The Horselenberg et al study took place in the Netherlands, so it also provides a cross-cultural replication. Thirty-four female psychology undergraduates participated in the research. The design was similar to Kassin and Kiechel, but the "SHIFT" key was the forbidden one, and there was no control group (no-witness conditions) or changes in pace of reading out letters.

Twenty-seven participants (82%) signed the confession (compliance), fourteen (42%) admitted to a confederate that they pressed the key (internalisation), and nineteen (58%) recalled doing it (confabulation). These results are higher than the figures found by Kassin and Kiechel (1996) (figure 1.5) ⁶.

⁶ The difference may be due to design and procedure variations between the two studies or the sample used (eg: only female students in Horselenberg et al).

Eight months previously the students had completed personality questionnaires and a measure of suggestibility. There was no difference in these scores between individuals who falsely confessed and those who did not ⁷.

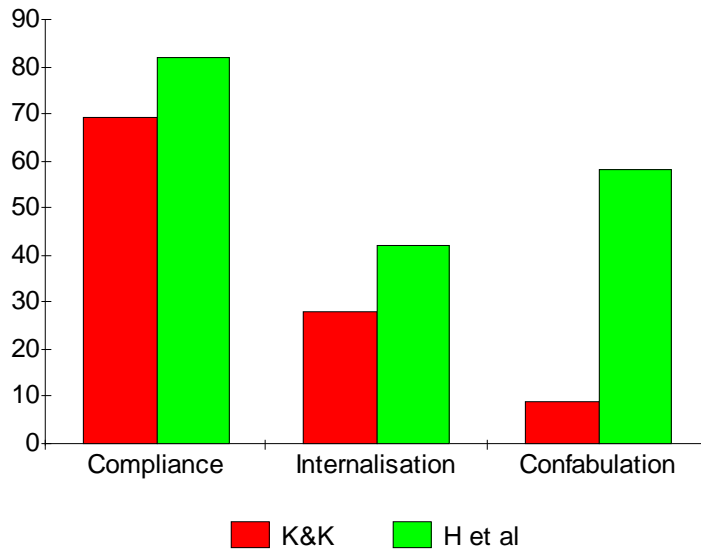


Figure 1.5 - Comparison of findings (%) between Kassin and Kiechel (1996) and Horselenberg et al (2003).

2. Ethical issues

The experiment involved a number of ethical issues:

i) The stress caused to the participants by the task, and, more importantly, when they were falsely accused of something they did not do. The relief expressed during debriefing shows that, for many participants, this was an issue.

ii) Deception was used by the experimenters in terms of the false witness, and the use of confederates.

Kelman (1967) noted that the use of deception will backfire on psychology in three ways:

- Participants will never believe experimenters, even when the experimenters are being truthful;
- Psychology generally cannot be believed;
- Research becomes a game of participant versus

⁷ Previous research found that real-life false confessors had higher suggestibility scores, for example (eg: Gudjonsson 1991).

experimenter: who is telling the truth?

But Gale (1994) suggested that studies "seem to indicate that student populations are less concerned about the ethicality of some aspects of psychological research than are psychologists or learned societies" (pp1169-1170).

Russano et al (2005), concerned about falsely accusing their participants of cheating, administered a follow-up questionnaire 3-10 weeks after the experiment. The mean for use of deception as justified was 5.26 (where 1 = not at all justified, 7 = very justified), and the stress reported was 4.27 (where 7 = extreme stress). The overall feedback was positive.

iii) A debriefing after the experiment may not have been enough to deal with the effects of the experiment.

There are two main aspects of any debriefing (Brewer 2001):

- Dehoaxing: the correction of fraudulent information used;
- Desensitizing: help participants deal with new knowledge of the self gained through the experiment, especially if it is negative.

However, there is the existence of the perseverance effect which challenges the effectiveness of dehoaxing. This effect is that first impressions remain despite information contrary later.

For example, participants are told that they scored badly on an IQ test in the experiment, in order to see their reaction to other tests. In the debriefing, they are informed that the IQ test scores were average. With the perseverance effect, there is a tendency is to believe that they scored badly on the IQ test as the truth.

An experiment by Ring et al (1970) has questioned the effectiveness of general debriefings. Fifty-seven undergraduates were used in an obedience experiment that involved making a loud noise in a victim's ear, and afterwards were given one of three debriefings. One group were not debriefed at all, the second group were given the traditional factual debriefing, and the last group received a traditional factual debriefing and explanations for their obedience behaviour. The first two groups were equally upset by the experiment when interviewed later, and 70% of them were suspicious about other experiments.

Thus a simple factual debriefing may not be enough when the participants have been involved in a very

emotionally upsetting experiment or where they have behaved unexpectedly for themselves.

Reynolds (1982) held up three criteria for ethics in research:

- The "utilitarian, cost-benefit criteria" - Is more learnt about human behaviour in this research, than if it was carried out another way, and if so, is it worth it? In other words, how important are the findings compared to the experiences of the participants?
- The outcome to the participants - What is the effect of the research upon the participants and their life?
- The integrity of the experimenter - Can the researcher maintain their integrity knowing how the results were collected?

In terms of the Kassin and Kiechel experiment, the authors would argue that the findings were important to understanding false confessions in real-life (with far-reaching consequences), and the experiment was not that unpleasant for the participants. In other words, the end justified the means.

1.6. REFERENCES

Brewer, K (2001) Ethical Issues for the Study of Human Behaviour by Psychologists Orsett, Essex: Orsett Psychological Services

Drizon, S.A & Colgan, B.A (2004) Tales from the juvenile confessions front. In Lasiter, G.D (ed) Interrogations, Confessions, and Entrapment New York: Kluwer Academic

Gale, A (1994) Ethical issues in psychological research. In Colman, A (ed) Companion Encyclopaedia of Psychology London: Routledge

Gudjonsson, G.H (1991) The effects of intelligence and memory on group differences in suggestibility and compliance Personality and Individual Differences 12, 503-505

Gudjonsson, G.H (2003) Psychology brings justice. The science of forensic psychology Criminal Behaviour and Mental Health 13, 159-167

Gudjonsson, G.H & MacKeith, J.A.C (1982) False confession, psychological effects of interrogation: A discussion paper. In Trankell, A (ed) Reconstructing the Past: The Role of Psychologists in Criminal Trials Deventer, Holland: Kluwer

Gudjonsson, G.H et al (2008) Personal versus impersonal relationship compliance and their relationship with personality Journal of Forensic Psychiatry and Psychology 19, 502-516

Gudjonsson, G.H et al (2009a) False confessions among 15- and 16-year-olds in compulsory education and the relationship with adverse life events Journal of Forensic Psychiatry and Psychology 20, 6, 950-963

Gudjonsson, G.H et al (2009b) Interrogations and false confessions among adolescents in seven countries in Europe. What background and

psychological factors best discriminate between false confessors and non-false confessors? Psychology, Crime and Law 15, 8, 711-728

Horselenberg, R et al (2003) Individual differences and false confessions: A conceptual replication of Kassin and Kiechel (1996) Psychology, Crime and Law 9, 1-8

Kassin, S.M (2008) False confessions: Causes, consequences, and implications for reform Current Directions in Psychological Science 17, 4, 249-253

Kassin, S.M & Kiechel, K.L (1996) The social psychology of false confessions: Compliance, internalisation, and confabulation Psychological Science 7, 125-128

Kassin, S.M & Norwick, R.J (2004) Why suspects waive their Miranda rights: The power of innocence Law and Human Behaviour 28, 211-221

Kassin, S.M & Sukel, H (1997) Coerced confessions and the jury: An experimental test of the "harmless error" rule Law and Human Behaviour 21, 27-46

Kassin, S.M & Wrightsman, L.S (1985) Confession evidence. In Kassin, S.M & Wrightsman, L.S (eds) Psychology of Evidence and Trial Procedure London: Sage

Kassin, S.M et al (2007) Police interviewing and interrogation: A self-reported survey of police practices and beliefs Law and Human Behaviour 31, 381-400

Kelman, H.C (1967) Human use of human subjects: the problem of deception in social psychology experiments Psychological Bulletin 67, 1-11

Reynolds, P.D (1982) Moral judgments: strategies for analysis with application to covert participation observation. In Bulmer, M (ed) Social Research Ethics London: Macmillan

Ring, K; Wallston, K & Corey, M (1970) Mode of debriefing as a factor affecting subjective reaction to a Milgram-type obedience experiment: an ethical inquiry Representative Research in Social Psychology 1, 67-88

Russano, M.B et al (2005) Investigating true and false confessions within a novel experimental paradigm Psychological Science 16, 481-486

2. ACCURATELY IDENTIFYING THE FACES OF CRIMINALS

- 2.1. Introduction
- 2.2. Improving facial composites
- 2.3. Appendix 2A - Frowd et al (2005)
- 2.4. References

2.1. INTRODUCTION

In court jurors place great emphasis on eye-witness testimony, particularly from confident eye-witnesses. Such eye-witnesses can be accurate, but they can also be wrong. For example, of 239 convictions overturned after the introduction of DNA testing, 73% were based on eye-witness testimony (Innocence Project quoted in Arkowitz and Lilienfeld 2010). The witness's ability to recall the face of the suspect using a facial composite or to identify the individual from mugshots is key.

Facial composite systems aid in the production of a visual likeness of a criminal suspect. After a crime, the construction of a facial composite of the suspect from witnesses' verbal recall of features is important to the police. But there are many problems with constructing faces from memory. One being the issue of whether faces are stored holistically or as individual features.

The police have traditionally made use of forensic artists, who draw the face as described by the witness, or manual facial composites that combine features. The original idea was based on a large number of pieces containing features of the faces that could be combined. Identikit, in the USA, used 470 features on acetate strips, while Photofit, in the UK, contained 855 features on jigsaw-like pieces (Frowd et al 2005).

Frowd et al (2005) summarised the main criticisms of manual composite systems:

- Low success rate in identification from a selection of faces (eg: 12.5%).
- Little difference in identification success rate between target face present or absent. It is expected that faces would be recognised better from composites when the target face is present.
- Edges of pieces of features create lines across the face which interfere with identification.
- Limited choice of facial features.

- Ignores the holistic nature of face perception.

Software packages, like E-FIT and PRO-Fit, have been subsequently developed. Computerised systems overcome some of these problems, and show greater success in experiments. But many experiments are artificial with the target face in-view as the composite is created or unrealistically short time between seeing the face and recall (Frowd et al 2005).

However, Davies et al (2000) found no difference in accuracy of composites between E-FIT and Photofit after one-minute exposure to a face in a laboratory experiment. While Frowd et al (2005) found E-FIT significantly better than Photofit (mean 12.5% vs 2.5% accuracy). Other techniques like EvoFIT produce whole faces which are adapted until a likeness is produced⁸.

In the experimental evaluation of these techniques (eg: Frowd et al 2005; appendix 2A) participants are shown an unfamiliar face, and later asked to describe it (ie: produce a facial composite). Then individuals who know the original face are asked if they recognise the facial composite. Accuracy using E-FIT and Pro-FIT with the "witnesses" soon after seeing the face (ie: two hours) is about 20%, but less than 10% if there is a real-life gap between seeing and recall (eg: two days). For artist-composites accuracy is about 10% with or without a delay (Frowd et al 2008a).

2.2. IMPROVING FACIAL COMPOSITES

Frowd et al (2008a) reported work over ten years at the Universities of Stirling and Edinburgh in Scotland and Central Lancashire in England to improve this process and the accuracy of facial composites.

1. Improving the recall process

Traditionally witnesses are asked to describe the face to the artist or police officer. But research has found that this process is improved if witnesses are asked to make personality judgments about the suspect beforehand (eg: intelligence, friendliness). Frowd et al (2008b) found that cognitive interviews which included making seven personality judgments about the suspect ("holistic cognitive interview"; H-CI) produced better facial composites than just the cognitive interview (CI) (table 2.1). It is felt that recall is improved because

⁸ See example at <http://www.uclan.ac.uk/scitech/psychology/research/evofit/overview.php>.

information about the physical aspects of the face are stored with judgments about the face/person, and thinking about the latter helps to trigger recall of the former.

Frowd et al (2008b) recruited 24 students from the University of Stirling, Scotland who reported not watching "EastEnders" (popular soap opera on BBC TV), and showed them a clip of that programme containing the target face (for 15-45 seconds on screen). After a wait of 3-4 hours, these "witnesses" constructed facial composites using PRO-fit computer software which builds a picture of the target face based on individual features.

Half of the participants did this in the context of the usual CI and half as part of the H-CI. In the latter, the participants were asked to judge the target face, from memory using a three-point scale (low/medium/high), on intelligence, friendliness, kindness, selfishness, arrogance, distinctiveness, and aggressiveness.

Then 36 regular viewers of "EastEnders" ("judges") were asked to name the characters from the facial composites produced by the "witnesses". The composites produced in the H-CI were named by 41% of "judges" compared to 9% for composites produced in the CI. The difference was even stronger for female composites than male ones.

Next fifty-four more regular viewers of "EastEnders" ("sorters") were asked to match the composites with target photographs. Either the whole face, internal feature (eg: nose, eyes, mouth) or external feature (eg: face shape, hair) composites were used in an independent design. Composites produced in the H-CI were matched more often than those from the CI (38% vs 23%). This difference held for whole face, internal and external feature composites.

"Witnesses" see target face briefly



3-4 hours later produced facial composite in CI or H-CI

(independent variable = type of interview)



"Judges" name target face from facial composite



"Sorters" match composites to target photographs

(dependent variable = accuracy of composite to target face)

Table 2.1 - Details of Frowd et al (2008b).

2. Improving the technology

Research has shown that familiar and unfamiliar faces are processed differently. With familiar faces the internal features (eg: eyes, nose, mouth) are more important, while external faces (eg: hair, face shape) matter more for unfamiliar faces. So individuals are more sensitive to changes in hair, for example, with strangers

(ie: fail to identify).

But witnesses are being asked to recall an unfamiliar face in most crime situations. Frowd et al (2008a) reported some success in overcoming this a "blurring technique". On the computer screen the hairstyle is blurred while the internal features are selected. This makes the internal features appear more prominent while maintaining the complete face ⁹.

3. Recall versus recognition

The traditional techniques are based on recall of the face, whereas Frowd et al (2008a) reported the developed of recognition-based techniques. For example, a witness is presented with 18 whole faces (with random features) on a computer screen, and are asked which one(s) are similar to the suspect. The selected ones are combined ("bred"), and the process continues until a likeness is produced.

There is a risk of interfering with the witness's memory if too many faces are presented. So ten scales (eg: age, facial weight, masculinity, honesty, threatening) were developed and the witness can change the computer image along these dimensions ¹⁰. This avoided too much inappropriate information at the beginning of the process.

Accuracy with EvoFIT has improved from 5% to 25% after a two-day delay by adding blurring and the use of dimensions to adjust the face (Frowd et al 2008a).

4. Multiple witnesses

If there are multiple witnesses to a crime, each one will produce a slightly different facial composite. This is a problem in terms of which one is correct.

Researchers have tried morphing these different composites with some success (eg: Bruce et al 2002) as "the consistent parts of the individual composites tend to be reinforced, and errors averaged out" (Frowd et al 2008a p671) ¹¹.

⁹ See example at http://www.thepsychologist.org.uk/archive/archive_home.cfm/volumeID_21-editionID_163-ArticleID_1381-getfile_getPDF/thepsychologist%5C0808frowd.pdf.

¹⁰ See example at http://www.thepsychologist.org.uk/archive/archive_home.cfm/volumeID_21-editionID_163-ArticleID_1381-getfile_getPDF/thepsychologist%5C0808frowd.pdf.

¹¹ See example at http://www.thepsychologist.org.uk/archive/archive_home.cfm/volumeID_21-editionID_163-ArticleID_1381-getfile_getPDF/thepsychologist%5C0808frowd.pdf.

5. Alternative views of the face

Most facial composites are front-view, but unfamiliar faces may be better processed in half profile (Bruce et al 1987). Ness et al (2001 quoted in Frowd et al 2008a) found improvements in accuracy when "witnesses" were asked to produce front-view and three-quarter views which were shown at the same time to those guessing the identity.

2.3. APPENDIX 2A - FROWD ET AL (2005)

Frowd et al (2005) compared five systems of facial composite used in the UK - Photofit (manual - focus on features), E-FIT and PROfit (computerised - features), Sketch (manual - holistic), and EvoFIT (computerised - holistic).

In their experiment there were ten conditions - five systems and whether the target face had low or high distinctiveness. Twenty-one volunteers rated on a scale of 1-7 whether the face stood out in a crowd (7) (high distinctiveness) or blended in with the crowd (1). Ten male celebrity targets were used¹².

Fifty volunteers were recruited as "witnesses" with ages ranging from 18 to 81 years old. Each participant was randomly shown a picture of a target face, and if they did not recognise the celebrity, they were asked to inspect the face for one minute. Then 3-4 hours later, the participants were asked to produce a facial composite based on one of the five systems in the context of a Cognitive Interview.

The accuracy of the facial composites were judged by 130 other participants' ability to recognise the celebrity from all composites from one system. The overall mean for recognition was 10.6%, but the accuracy varied with distinctiveness of face and system used. Highly distinctive faces were recognised three times more often than low distinctive ones on average, and always better with each system. The best systems were E-FIT (mean 19.0%) and PROfit (17.0%), followed by Sketch (9.2%), Photofit (6.2%) and EvoFIT (1.5%) (figures 2.1 and 2.2)¹³.

¹² Low distinctiveness - Michael Owen (footballer), Damon Albarn and Stephen Gateley (singers), Craig Phillips and Noah Wyle (actors). High distinctiveness - Robbie Williams and Noel Gallagher (singers), Brad Pitt (actor), Andre Agassi (tennis player), David Beckham (footballer).

¹³ Accuracy of recognition was measured by the conditional naming rate - calculated by dividing the number of correctly named composites by the number of correctly named targets and averaged over participants (Frowd et al 2005 p44).

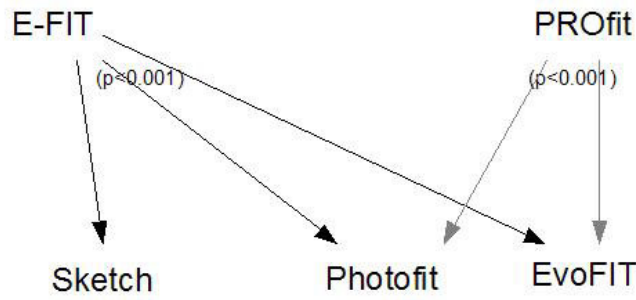


Figure 2.1 - Significant differences in accuracy between the five systems.

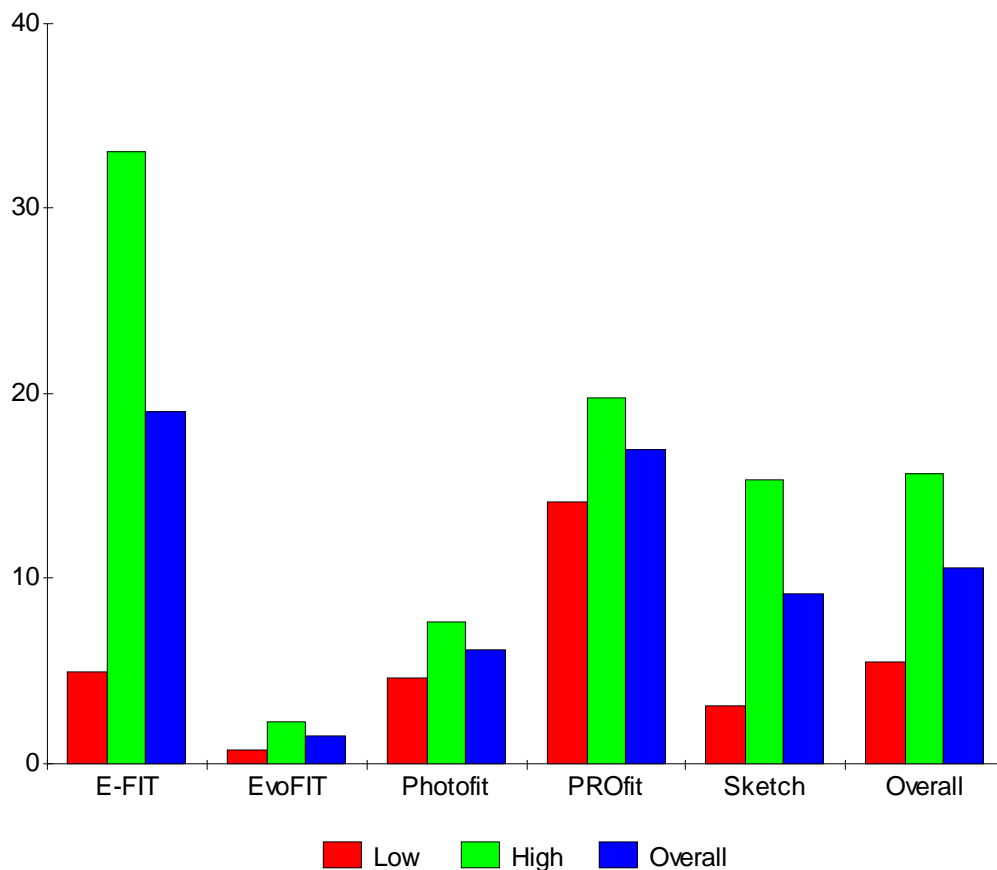


Figure 2.2 - Accuracy (%) of recognition of composites based on system and distinctiveness.

The accuracy of composites was also measured by asking fourteen undergraduates to sort all the composites with the target faces present. The overall accuracy was 65% with E-FIT, PROfit and Sketch composites nearer 75%, and Photofit and EvoFIT nearer 50% (figure 2.3).

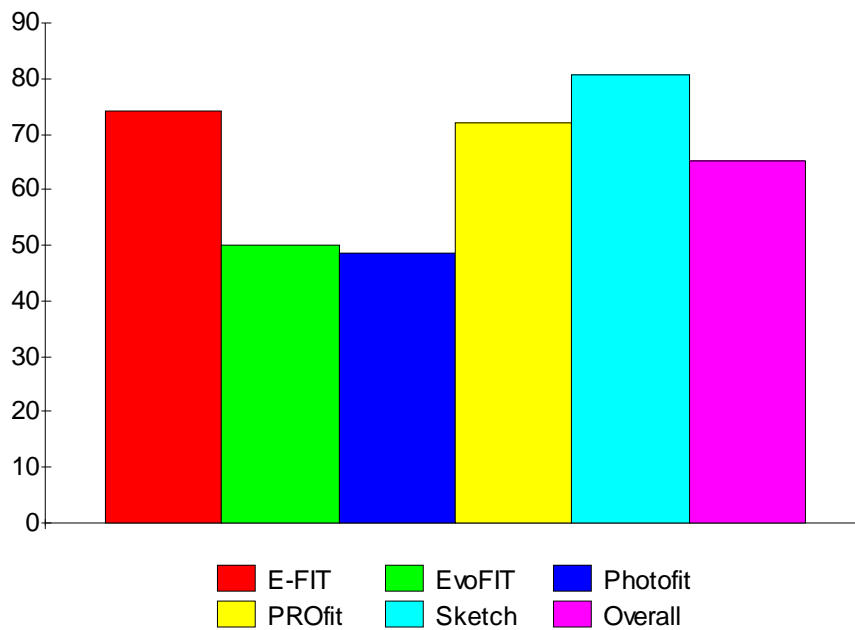


Figure 2.3 - Accuracy (%) of sorting of composites.

Figure 2.4 summarises the stages of the experiment.

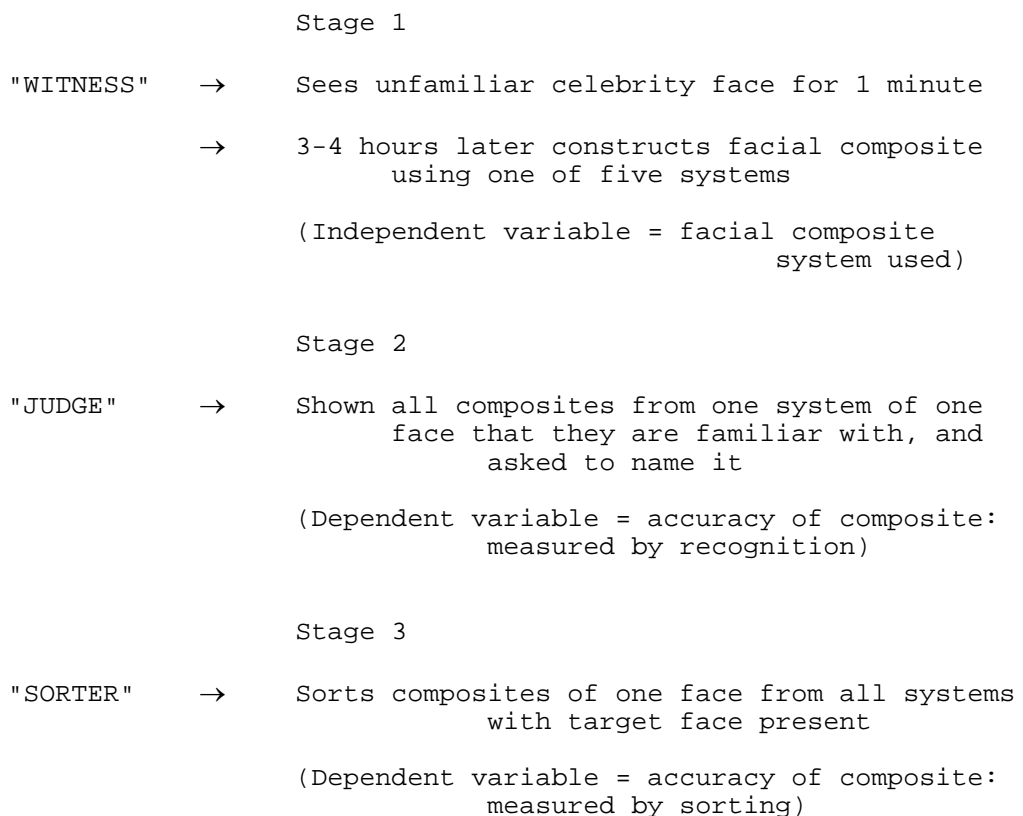


Figure 2.4 - Stages of experiment by Frowd et al (2005).

Methodological Issues

1. The choice of ten male celebrity faces from the Internet. Why no female faces? Frowd et al argued that "since most crimes are committed by young males..., these stimuli appear appropriate" (p38).

Full-face photographs with a neutral expression were selected. Young faces were used to limit ageing effects (eg: crows feet, forehead wrinkles and under-eye bags), "features that may exaggerate operator differences" (p38).

2. The length of delay between seeing the face and construction of composite was shorter (3-4 hours) than in real-life (at least 24-36 hours). The authors explained: "A practical solution was chosen for operators and witnesses in this study with a 3-4-hour target delay, enabling targets to be presented in the morning and composites constructed in the afternoon: a tightly controlled procedure that engaged witnesses for a day" (p40). Though researchers would like to control for all extraneous variables, there are also practical considerations like time and money involved.

3. Experienced operators were used for each system, and each operator worked on the same target only once. They were blind to the target face being constructed by the "witness".

The operators did not construct more than one composite of the same target, and so avoided bias from previous composites, but different operators in each condition might have influenced the composite quality. Frowd et al concluded: "Our design is thus a compromise - avoiding operator contamination at the expense of potentially elevating operator differences" (p49).

4. For the "witnesses", participants were recruited who had not constructed composite before. They were students, staff, and others from and around the area of the Universities of Stirling and Abertay in Scotland, probably (as no information is given). There were twenty males and thirty females. They were randomly allocated to one system of composite.

Overall, the mean age was forty years old, but the mean age of the Sketch condition was 56 years and 32 years for the Photofit condition. Random allocation to independent groups cannot guarantee that the groups are equal on demographic variables. Matching of participants beforehand would remove this problem. The authors noted that statistical analysis of the groups based on age

found no significant differences, which suggested that it was not a problem.

5. The participants selected a photograph from an envelope at random, and said if they recognised the person. If they did, they selected another photograph. If the participant recognised all the ten photographs, they did not take part in the study.

This key part of the experiment (ie: witness) was based on the word of the participants, and their "random" choice of photograph. It is not necessarily that the participants would lie about recognising the celebrity, but that during the wait to construct the facial composite, recognition may occur (even at a below-conscious level). Frowd et al admitted: "Clearly, as all our targets are famous, witnesses may have indeed seen them in the past, though not sufficiently often for recognition to have occurred during our study. Thus, it would appear more appropriate to say that our witnesses claimed not to be consciously familiar with their target" (p49). Unfamiliar faces of individuals not famous would be desirable.

6. Using experienced operators to construct the composite was realistic. A procedure as recommended by the Association of Chief Police Officers was used. This began with "rapport-building" (informal chat), then the Cognitive Interview. This involved free recall of the face without interruption followed by cued recall of features based on what had been said earlier.

Then the specific phase of the system occurred. Though the first part of the process was standardised, this phase varied between systems used and the individual operator. It could not be standardised, and this leaves the possibility for extraneous variables, like personality of operator, ease of use of system, and time for construction of composite.

E-FIT and PROfit involved the combination of features on the computer screen which were adapted (exchanged, resized or re-positioned) in response to the "witness". These systems took on average 60-70 minutes to complete. Photofit used jigsaw pieces containing features, and lasted an average of 45 minutes. The sketch artist drew features selected by the witness, and then fleshed out the detail (lasting an average of 2 hours and 15 minutes). With EvoFIT whole faces were presented and bred together until the "best-face" was produced after 2½ hours on average.

The operators were also not equal because of the

skill required as a sketch artist, for example, is different to that of operators of the other systems. Frowd et al reported that "an informal analysis (by the first author) noted that sketches created in the study tended to contain more detail for the face shape, hair, eyes, eyebrows and mouth; and less detail for the forehead, cheeks, chin and areas around and including the nose" (p47). Furthermore, "our sketch artist has informally attempted to add extra artwork to make a sketch appear more lifelike. When this was carried out, witnesses were dissatisfied and the extra artwork had to be removed" (p47). While the operator of EvoFIT experienced with the software, despite being trained, because "the system was rather complex to operate" (p48).

7. The participants, acting as judges, who recognised the celebrity from the composites had to name them, or, if not, give clear "tip of tongue" information about them - eg: "Footballer, used to have a Mohican hairstyle, married to Posh Spice" (pp42-43) was acceptable for David Beckham, but not just "footballer".

The participants were shown the target photographs to name after this process. The composite naming score showed the number of composites recognised in relation to the celebrities familiar with. Familiarity with the celebrities varied between 54-100% (ie: some celebrities were only known by half the participants and some by all).

8. The rating of the distinctiveness of a face was based upon the judgments of 21 volunteers: "They were told that they would be shown a number of photographs of celebrities, asked to imagine meeting each person at a railway station in amongst their peers (young, white males) and to rate these on distinctiveness from 1 to 7 (1 = average, blend in to the crowd and 7 = very distinctive, stand out from the crowd)" (p39).

This process reduced 22 faces to ten. The low distinctiveness faces chosen for the experiment had mean scores below four, and the high distinctiveness faces above four (figure 2.5).

AVERAGE, BLEND IN TO THE CROWD (1)

↑

2.6	Michael Owen
2.9	Damon Albarn
3.3	Stephen Gateley
3.6	Craig Phillips
3.7	Noah Wyle

↑

Low distinctiveness

High distinctiveness

↓

4.3	Robbie Williams
4.9	Brad Pitt
5.0	Andre Agassi
5.1	David Beckham
5.8	Noel Gallagher

↓

VERY DISTINCTIVE, STAND OUT FROM THE CROWD (7)

Figure 2.5 - Mean distinctiveness ratings for each celebrity's face.

9. In terms of ethical issues related to the study of human participants:

- Volunteer participants recruited which deals with issue of forced participation by psychology students, for example, as part of their course requirements.
- Experiment not stressful for participants as it was emphasised that they were "passive witnesses".
- The "witnesses" were paid £10 for their participation. There is discussion over the appropriateness of payment of participants, but when the amount is small it covers travel expenses only and there is less of a concern. But does it change the nature of the relationship between participant and experimenter?
- The "witnesses" were debriefed after the construction of the facial composites with as much detail of the project as relevant.
- The anonymity of the participants was maintained in terms of their names etc, but also no details of where experiments took place.

2.4. REFERENCES

Arkowitz, H & Lilienfeld, S.O (2010) Do the "eyes" have it? Scientific American Mind January/February, 68-69

Bruce, V et al (1987) The basis of the 3/4 view advantage in face recognition Applied Cognitive Psychology 1, 109-120

Bruce, V et al (2002) Four heads are better than one. Combining face composites yields improvements in face likeness Journal of Applied Psychology 87, 894-902

Davies, G.M et al (2000) Facial composite production: A comparison of mechanical and computer-driven systems Journal of Applied Psychology 85, 1, 119-124

Frowd, C.D et al (2005) A forensically valid comparison of facial composite systems Psychology, Crime and Law 11, 33-52

Frowd, C.D et al (2008a) Changing the face of criminal identification Psychologist 21, 8, 668-672

Frowd, C.D et al (2008b) Improving the quality of facial composites using a holistic cognitive interview Journal of Experimental Psychology: Applied 14, 3, 276-287

Ness, H et al (2001) Are two views better than one? Paper presented at the 11th European Conference on Psychology and Law, Lisbon, June 2001

3. MEASURING MALINGERING IN THE CRIMINAL JUSTICE SYSTEM

- 3.1. Measuring and defining malingering
- 3.2. Detection of malingering
 - 3.2.1. Experimental simulation
- 3.3. References

3.1. MEASURING AND DEFINING MALINGERING

Honesty and lying are key issues in forensic and criminal psychology, particularly in relation to detecting them accurately. They are usually linked to what an individual says (eg: a suspect's confession). But there is another area related to what a person does, particularly in relation to the "symptoms" they show. This is a form of pretending known as malingering (or symptom exaggeration). For example, an exaggeration of the symptoms in personal injury civil cases (which could be as high as one-third; Sullivan et al 2007).

Merckelbach et al (2009) noted that malingering could be viewed as "a rational strategy among people who face a difficult situation and who try to serve their own interests in the best possible way" (p379-380). It can occur in any situation where it benefits the individual to use such a strategy.

In DSM-IV malingering is defined as the "intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs" (quoted in Jaffe and Sharma 1998 p552) ¹⁴.

Mittenberg et al (2002) performed an anonymous postal questionnaire study with members of the American Board of Clinical Neuropsychology, of which 144 neuropsychologists completed the survey. These practitioners produced assessments in personal injury, disability, criminal, and medical cases. Mean estimates

¹⁴ Jaffe and Sharma (1998) listed six reasons for malingering - "to avoid punishment, responsibility or execution for a crime, or to render them worthy of mitigation of an assigned penalty"; to avoid military service or hazardous duty; financial gain; "to facilitate transfer from prison to a hospital, from which a defendant may hope to escape, do easier time, or take advantage of the mentally ill inmates"; "to gain admission to a hospital for free room and board, or to avoid police apprehension"; to obtain drugs to get high. The authors noted that "numerous patients present at the walk-in clinic claiming to have lost their Ativan (benzodiazepine) and need more; I cannot recall one patient ever presenting to the walk-in clinic who claimed they lost their anti-psychotic medication" (Jaffe and Sharma 1998 p552).

were calculated for the number of cases where malingering was believed to be occurring (table 3.1).

Personal injury cases	29
Disability or worker compensation cases	30
Criminal cases	19
Medical or psychiatric cases	8

Table 3.1 - Mean rates (%) of malingering in different cases.

Mild head injury was most often viewed by the neuropsychologists as exaggerated (39% of cases), followed by "chronic fatigue" (35%) and chronic pain (31%).

These figures were based on the opinions of the neuropsychologists, but what evidence did they have for these views? Part of the questionnaire asked about the basis of the opinion. The most common evidence was "severity of cognitive impairment inconsistent with condition" (65% of respondents), "pattern of cognitive test performance inconsistent with condition" (65%), "scores below empirical cut-offs on forced choice tests" (57%), and "discrepancies among records, self-report, and observed behaviour" (56%). Most neuropsychologists used multiple indicators to arrive at their opinion (7 out of nine indicators).

Higher rates of symptom exaggeration were reported for individuals referred by defence attorneys in civil cases, and by prosecutors in criminal cases.

In 1994, the state of California in the USA enacted a law called "three strikes and you're out" where the third criminal offence leads to a minimum of twenty-five years if convicted (irrelevant of the actual crime)¹⁵. In this situation, defendants are motivated to show psychiatric problems, and thereby avoid such punishment. In other words, extreme sentences for minor crimes encourages malingering.

Jaffe and Sharma (1998) reported assessing nine male defendants in this situation. Of which eight of them were believed to be claiming false psychiatric symptoms, and were rated as competent to stand trial. All of them volunteered information about their symptoms which included amnesia for certain information (eg: their name, where they were), visual hallucinations, and bizarre behaviours (eg: eating bugs).

¹⁵ The first two convictions must be for serious crimes like possession of narcotics, burglary and violent crime, but the third offence can be petty and the rule still applies.

3.2. DETECTION OF MALINGERING

Jaffe and Sharma (1998) noted that "contrary to intuition and popular myth, facial expression and eye contact are poor indicators of truthfulness" (p552). They recommended the following pointers:

- Withholding of information by the individual based on the belief that less information given to the examiner is better.
- Lack of co-operation with the examiner through fear that a rigorous examination to find them out.
- Exaggeration of symptoms. "Malingerers often mistakenly believe that the more bizarre they appear, the more convincing they become" (p553).
- Inconsistency in behaviour (eg: acts normally when thinks not being observed).
- Symptoms inconsistent with mental disorder being claimed. For example, amnesia about own name, but able to remember other facts. Such amnesia would be a sign of dementia and would not be selective.
- Deception as general pattern of behaviour (eg: criminal fraud).
- Malingering is difficult to maintain over a long period of time.

These are useful pointers, but how to detect malingering (or "presentation of non-credible symptoms"; Boone 2007) accurately and reliably? This is usually done with psychometric tests, like the Structured Inventory of Malingered Symptomatology (SIMS) (Smith and Burger 1997) which seeks to detect over-endorsement of symptoms. The design of such tests is crucial.

3.2.1. Experimental Simulation

One technique used in the design of psychometric tests is the experimental simulation study. A group of individuals, usually students,¹⁶ are asked to take the test, with half pretending to have the condition under scrutiny (fakers) and the other half answering honestly.

¹⁶ Greve and Bianchini (2004) believed that simulation samples should be similar to the usual demographic characteristics of malingerers.

The fakers are given instructions such as: "Assume that you are in a situation where it would benefit you greatly to appear mentally retarded. Therefore, please respond so that you present yourself as someone with serious psychological problems" (Holden et al 2003 p1110 quoted in Merckelbach et al 2009 p379). A test has discriminability will distinguish between the two groups by their scores.

Merckelbach et al (2009) felt that this placed a great emphasis on the instructions given to participants in experimental simulations, and they set out to test the effects of different scenarios.

For their study, 486 first-year students in medicine, health sciences, and psychology in the Netherlands completed the Wildman Symptom Checklist (WSC) (Wildman and Wildman 1999). This is a self-reported scale to detect over-endorsement of symptoms and positive personal qualities. Three sets of ten malingering items are embedded within the total of sixty items including ordinary ones like "I frequently experience headaches".

- Non-credible physical symptoms - eg: "The buzzing in my ears keeps switching from the left to right", "I have allergies that I suffer from only at night".
- Non-credible mental symptoms - eg: "I have a terrible fear of street signs", "Someone is plotting to kill me and I know they will strike at midnight".
- "Fake good" tendencies - eg: "I never make a good story better", "I never find it difficult to talk to strangers".

Each item is scored as "yes" or "no" giving a range of 0-30 for malingering (with a total score of four as the cut-off point; Wildman and Wildman 1999).

The participants were divided into four groups to fill out the WSC:

i) The control group (n = 387) who were asked to complete the WSC honestly.

ii) The "manslaughter group" (n = 33) read about a real Dutch case of a young man who, while trespassing, accidentally knocked some stones off a wall that hit a young girl and killed her. Participants were asked to fake a serious psychological condition to minimise criminal responsibility in this situation.

iii) The "wrongful death group" (n = 33) read about a real Dutch case of a railway worker who failed to signal an incoming train and another worker was killed.

Participants were asked to fake mental retardation, so that the railway company would be held responsible for delegating safety measures to a person with insufficient cognitive capabilities.

iv) The "personal injury group" (n = 33) read about a real-life case of a worker in a factory producing paint coatings who filed a civil suit against the company for chronic toxic encephalopathy (brain damage from toxic fumes). Participants were asked to fake the symptoms of this condition.

The ability of the WSC scores to distinguish honest replies from fakers was the key focus. The malingering of the personal injury group was harder to spot than the other two faker groups. "In other words, the civil case vignette produced a less intense and therefore more difficult to detect form of malingering than the criminal case vignettes" (Merckelbach et al 2009 p383) (figure 3.1).

The authors concluded that malingering can be difficult to spot in different situations, particularly if a single test with a simple cut-off point is used. Malingering is a "strategic behaviour that sometimes takes the form of positive symptomatology (ie: bizarre symptoms) combined with feigned impairments (ie: pseudo-neurological dysfunctions) and at other times restricts itself to the latter type of responding" (pp384-385).

The accuracy of any test depends upon (Greve and Bianchini 2004):

a) Sensitivity - This is the "true positive" (hit) rate - ie: the number of person with the condition as correctly rated by the test divided by all persons with the condition. For example, if 100 out of 1000 people have the condition and the test correctly spots ninety of them, the hit rate is 90%. The ten people undetected are "false negatives" (ie: have condition and missed by test).

Sensitivity is reduced by the test being too obvious to the malingerer or by "attorney coaching", but increased by the use of multiple detection techniques.

b) Specificity - This is the "true negative" rate - ie: the number of persons without the condition who had a negative test result divided by all persons without the condition. Poor specificity produces "false positives" (individuals diagnosed with the condition who do not have it).

A cut-off score must balance sensitivity and specificity. The question is what is more serious - to



Figure 3.1 - Number of participants (%) in each group scoring four or more on the WSC (and mean score of each group out of 30).

miss true cases or to include false positives? Greve and Bianchini (2004) argued that tests of malingering with a single cut-off point should have a false positive rate of

zero (perfect specificity). It would be better to collect data and to set probability levels for the behaviour being fake.

3.3. REFERENCES

Boone, K.B (2007) A reconsideration of the Slick et al (1999) criteria for malingered neurocognitive dysfunction. In Boone, K.B (ed) Assessment of Feigned Cognitive Impairment: A Neuropsychological Perspective New York: Guilford

Greve, K.W & Bianchini, K.J (2004) Setting empirical cut-offs on psychometric indicators of negative response bias: A methodological commentary with recommendations Archives of Clinical Neuropsychology 19, 533-541

Holden, R.R et al (2003) Experimental faking in self-reported psychopathology: Uni-dimensional or multi-dimensional? Personality and Individual Differences 35, 1107-1117

Jaffe, M.E & Sharma, K.K (1998) Malingering uncommon psychiatric symptoms among defendants charged under California's "three strikes and you're out" law Journal of Forensic Science 43, 3, 549-555

Merckelbach, H et al (2009) Experimental simulation: Type of malingering scenario makes a difference Journal of Forensic Psychiatry and Psychology 3, 378-386

Mittenberg, W et al (2002) Base rates of malingering and symptom exaggeration Journal of Clinical and Experimental Neuropsychology 24, 8, 1094-1102

Smith, G.P & Burger, G.K (1997) Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology Journal of the American Academy of Psychiatry and the Law 25, 183-189

Sullivan, K et al (2007) Methods of detecting malingering and estimated symptom exaggeration base rates in Australia Journal of Forensic Neuropsychology 4, 49-70

Wildman, R.W & Wildman, R.W (1999) The detection of malingering Psychological Reports 84, 386-388

4. GENES AND CRIME, AND THE COMPLEXITY OF THE RELATIONSHIP

- 4.1. Stephen Mobley case
- 4.2. Genes and crime
- 4.3. References

4.1. STEPHEN MOBLEY CASE

In February 1991, Stephen Mobley shot the manager of a "Domino's Pizza" store in Oakwood, Georgia, USA during a robbery. He shot the manager (John Collins) in the back of the head without mercy. He was caught soon afterwards (Denno 2006).

Mobley came from an economically privileged home with no history of abuse, yet from an early age he exhibited behavioural problems including cheating, lying and stealing. By his mid-20s he was involved in many armed robberies including the "Domino's Pizza" in Oakwood. "While awaiting trial for Collins's death, Mobley's aggression was out of control: He fought continually with other inmates, sodomised his cellmate, tattooed the word 'Domino' on his own back, and verbally taunted and threatened prison guards. As a youth and as an adult, seemingly no amount of counselling or punishment could contain Mobley's outbursts" (Denno 2006 p215).

His defence attorneys found out from family members that there was a history of violence in the Mobley family - "The Mobleys were either behaviourally disturbed or business achievers, and, in a number of cases, they were both" (Denno 2006 p216). The Mobley family seemed similar to a case reported by Brunner et al (1993). This was a Dutch family where fourteen males over four generations showed borderline mental retardation and serious behavioural problems (impulsive aggression, arson, attempted rape, and exhibitionism).

The affected males showed a defect on the p11-p21 region of X chromosome leading to monoamine oxidase A (MAOA) deficiency, which was passed from mother to son. MAOA deficiency leads to low serotonin, which is linked to impulsive aggression, and to rapid eye movement (REM) sleep deprivation (which is also linked to aggression) ¹⁷.

The ideas in this work were used in a 1994 appeal against his death sentence by Mobley. It was argued that

¹⁷ Cases et al (1995) genetically engineered mice with the gene mutation observed by Brunner et al (1993), and they showed increased aggression in males.

he showed a "genetic propensity for misconduct" and could not help his behaviour. Thus this was a mitigating circumstance. The appeal failed and Mobley was executed in March 2005.

4.2. GENES AND CRIME

The Stephen Mobley case became influential in the debate about the genetics of crime. "A common stereotype is that an individual's 'genotype' or 'genetic constitution' is static, as though there is a 'crime gene' that 'hard-wires' certain people to violate the law. But this perspective, however entrenched in the public's mind, has no scientific support. Rather, an overwhelming amount of evidence shows that 'environments influence gene expression'" (Denno 2006 p213). So rather than a simple deterministic gene causing behaviour with no choice, genes produce tendencies which, depending on other factors, may or may not lead to certain behaviour (figure 4.1).

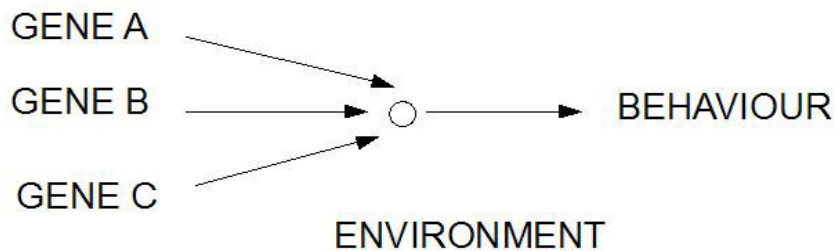
An alternative to these relationships in figure 4.1 could be to see the process as additive. It is a combination of factors, one on the other, that leads to the criminal behaviour. Many people will experience the factors individually, but only added together will they lead to the criminal behaviour. For example, in figure 4.2, hypothetically starting with 100 people who experience poverty, with the combination of factors there are a small number of individuals who end up as criminals. The number of factors will vary in different situations.

Leaving aside that one gene causes one behaviour, apart from specific conditions like cystic fibrosis, being wrong, the more knowledge that is gained about the human genome, the more complex it appears to be, "while the mechanisms of genetic inheritance now appear more elusive than ever" (Le Fanu 2010 p43).

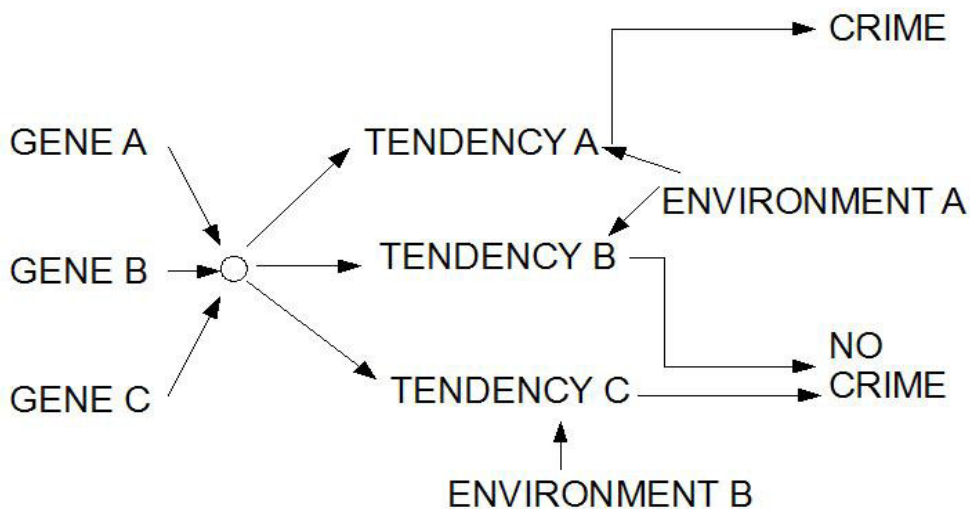
(1) Single gene causes behaviour
(irrelevant of environment)



(2) Multiple genes cause behaviour
(irrelevant of environment)



(3) Multiple genes interact to produce tendencies
or probabilities depending on the environment



(1) This relationship is too simple and unlikely to be relevant for explaining crime.

(2) It could be that a number of genes interact together to produce the behaviour. This is still very deterministic as the environment has no influence.

(3) This model allows for greater flexibility between genes, environment and behaviour. The genes produce tendencies or biases towards certain behaviours, and depending on the environment, certain behaviour occur. For example, a tendency towards aggression in one situation leads to an excellent sports player or business leader, while in another situation it manifests as violent crime.

Figure 4.1 - Three possible relationships between genes, environment and behaviour.

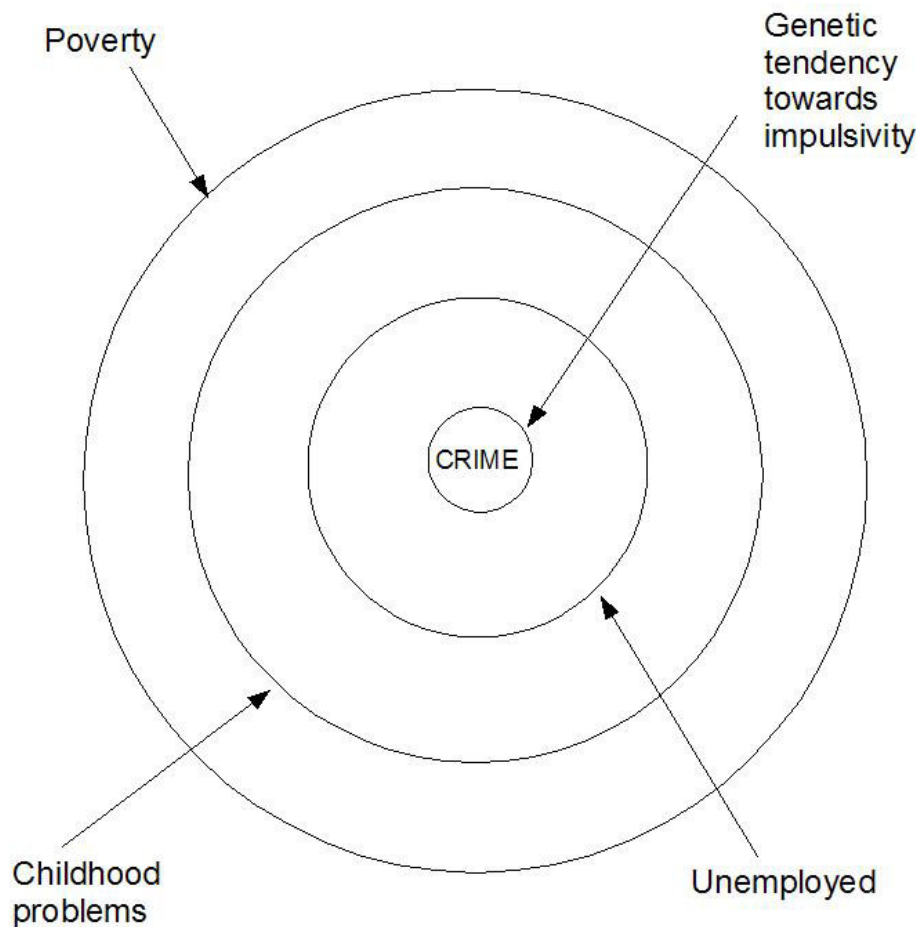


Figure 4.2 - Hypothetical example of an additive model to explain criminal behaviour.

4.3. REFERENCES

Brunner, H.G et al (1993) Abnormal behaviour associated with a point mutation in the structural gene for monoamine oxidase A Science 262, 578-580

Cases, O et al (1995) Aggressive behaviour and altered amounts of brain serotonin and norepinephrine in mice lacking MAOA Science 268, 1763-1768

Denno, D.W (2006) Revisiting the legal link between genetics and crime Law and Contemporary Problems 69, 209-257

Le Fanu, J (2010) The disappointments of the double helix: A master theory Journal of the Royal Society of Medicine 103, 43-45

5. FEMICIDE AS A PUBLIC HEALTH ISSUE

- 5.1. Introduction
- 5.2. Campbell et al (2003)
 - 5.2.1. Key methodological issues
- 5.3. Appendix 5A - Domestic abuse
- 5.4. References

5.1. INTRODUCTION

"Femicide" is a term for the murder or homicide of women. In the USA, it is the seventh leading cause of premature death for all women, and number one for African American women aged between 15-45 years old. The perpetrator is more often an "intimate partner" (husband, lover, ex-husband, ex-lover) (Campbell et al 2003). In England and Wales in 2005-6, one-third of the 248 total female homicide victims were killed by an intimate partner (and about a quarter of male victims) (BMA Board of Science 2007). Femicide is an extreme version of "domestic abuse" (appendix 5A).

Femicide can be viewed in the same way as any other public health issue, like smoking and HIV, by identifying risk factors with the possibility of early intervention. This creates the category of women "at risk", where prevention of the death is possible.

Public health or epidemiological studies collect large amounts of data, and, through statistical analysis, isolate the variables/factors that distinguish between high and low risk groups. One design of study is case-control, which compares cases with non-cases (control) for differences between the groups. It is a form of quasi-experimental design.

5.2. CAMPBELL ET AL (2003)

Using the case-control method, Campbell et al (2003) collected 220 cases from 1994 to 2000 in eleven cities in the USA where the female victim was killed by an intimate partner (current or former) (who was known to the police - ie: arrested/convicted). Details of the relationship between the victim and the perpetrator prior to the murder were ascertained from knowledgeable informants (eg: relative or friend of victim).

The control group were 343 abused women (18-50 years old) living the same cities. These were selected by stratified random-digit telephone dialling. "Abused" was defined as the women being "physically assaulted or threatened with a weapon by a current or former intimate

partner during the past two years" (p1089).

English- and/or Spanish-speaking interviewers¹⁸ questioned the case informants and the control women (usually by telephone) for 60-90 minutes. Information about their backgrounds and relationships were collected as well as questions from psychometric tests like Danger Assessment (Campbell 1995) (eg: "Does your partner try to control all of your daily activities?", "Have you ever been beaten by him when pregnant")¹⁹.

This type of study is not an experiment, so the standardisation of procedure will have flaws. For example, the use of a large number of different interviewers, though performing the same structured interviews, is an extraneous variable. This study is concerned with the patterns across large amounts of data rather than with complete control of situational variables as in the laboratory experiment (table 5.1). It is hoped that the large quantity of information will iron out any individual differences.

CASE-CONTROL METHOD	LABORATORY EXPERIMENT
<p>1. Quasi-experimental design which allows flexibility to study real-life events, but does not involve randomisation and control of confounding variables.</p> <p>2. Procedure is similar for each participant, but not entirely standardised and identical.</p> <p>3. The focus is upon the relationship between a large number of factors as established by statistical analysis.</p>	<p>1. Greater control over participants (eg: randomisation to conditions) and variables.</p> <p>2. Usually smaller scale which aids control, particularly as participants come to laboratory to be studied.</p> <p>3. Not possible to study certain topics like femicide with this method.</p>

Table 5.1 - Key differences between the case-control method as used by Campbell et al (2003) and laboratory experiments.

The results were analysed using logistic regression models²⁰. These estimate the independent associations

¹⁸ Non-English- and Spanish speakers were excluded from the study.

¹⁹ There are twenty items and each one is answered "yes" or "no" (see copy at <http://www.dangerassessment.org/WebApplication1/pages/da/DAEnglish2010.pdf>).

²⁰ Regression is an extension of a simple (bivariate) correlation. With a correlation, the relationship between two variables is established as positive (both variables increase or decrease together), negative (inverse relationship), or no correlation. Because the relationship is two-way (ie: variable A could cause variable B and vice versa), it is not possible to establish a definite causation. Where there are many variables, the relationship will be more complex. Regression shows the relationship between several causes/predictors (or independent variables), and with an outcome (dependent variable). There are

between variables (risk factors) and femicide. Statistical analysis showed the risk factors that independently increased or decreased the likelihood of femicide by the intimate partner (table 5.2).

PERPETRATOR	VICTIM
<p>Increase risk:</p> <ul style="list-style-type: none"> • Lack of employment 5.09 * (1) • Access to firearm 7.59 (2) • Use of illicit drugs 4.78 (2) • Highly controlling - & separated 8.98/ & together 2.90 (4) • Previous threats with weapon 4.08 (5) • Threats to kill victim 2.60 (5) • Abuser used gun 41.38 (7) <p>Decrease risk:</p> <ul style="list-style-type: none"> • College education 0.31 (1) • Previous arrest for domestic violence 0.34 (6) 	<p>Increase risk:</p> <ul style="list-style-type: none"> • Previously lived together, but separate at time of incidence 3.64 (3) • Left or asked abuser to leave 3.20 (3) • Victim had child by previous partner living in home 2.23 (3) • Victim left abuser for other person 4.91 (7) • Victim left abuser (for another reason) 4.04 (7) <p>Decrease risk:</p> <ul style="list-style-type: none"> • Higher education 0.31 (1) • Sole access to gun 0.22 (2) • Never lived together 0.39 (3)

(* = adjusted odds ratio ²¹; a figure less than one means reduced risk with that variable present)

- (1) Socio-demographic variables
- (2) General risk factors for homicide
- (3) Relationship variables
- (4) Abuser controlling victim's behaviour variables
- (5) Threatening behaviours by abuser
- (6) Previous physical abuse
- (7) Incidence variables (eg: triggers to murder)

Table 5.2 - Significant risk factors for intimate partner femicide.

different regression models, and they vary in details like the order in which to enter predictors into the equation. The three main models of regression are (Coakes and Steed 2003):

- Standard or simultaneous - All predictors entered together to see the relationship between all the causes and the outcome.
- Hierarchical - The predictors are entered in order of the theoretical basis of the study.
- Stepwise - The predictors are entered in different ways depending on statistical criteria.

²¹ Odds ratio (OR) = ratio of affected to unaffected individuals in one group divided by same ratio in another group (Petrie 1987). For example, fifty of 100 married are abused compared to ten of 100 unmarried women. The OR = $(50 \div 50) \div (10 \div 90)$; = $1 \div 0.11 = 9.09$. Thus the OR of being abused as married woman is nine times greater than as unmarried in this example.

Putting all the variables together, femicide risk was increased with unemployed abusers who were usually highly controlling and had previously threatened the victim with a weapon. The trigger was the victim leaving (for another person or not) and a gun was involved. The risk was reduced significantly if the couple had never lived together, and the abuser had been previously arrested for domestic violence.

The following conclusions were drawn from the study:

- There was previous physical abuse before the femicide (in the majority of cases).
- Arrest for this previous domestic violence was a protective factor.
- Unemployment of abuser was the only socio-demographic variable that distinguished cases from controls.
- Never living together reduced the risk, while living together with a child from the victim's previous relationship, and then leaving the abuser increased the risk (estrangement).

5.2.1. Key Methodological Issues

1. The reliance on informants for information about femicide cases compared to information gathered directly from the abused women. The authors felt that the informants' information was not deliberately inaccurate. However, there is the "honest liar" situation where individuals report what they believe to be the correct information, but it is wrong (eg: due to memory error).

Informants giving "don't know" responses were treated as absence of the behaviour, which would produce an underestimate. Ultimately, this would lead to a type II error - failure to reject the null hypothesis when the results were significant.

2. The information about the femicide cases and the informants to contact came from police records. These may have included inaccurate information about the relationship between the victim and the perpetrator. For example, how former is a former partner? Is it only the most recently ended relationship? Some perpetrators may have been from previous relationships in the past and not classed as intimate in police records.

3. The study did not include women who resided outside large urban areas, nor women in the control group without telephones.

4. The women in the control group who refused to participate may have been those with the highest risk of femicide. This would have inflated the estimates of risks associated with femicide.

5. The greatest strength of the study was a control group of physically abused women to compare to the femicide cases.

6. The analysis of large amounts of data for general patterns downplays the exceptions. Furthermore, it ignores the experience of the victims (which could only be collected for the live women). But qualitative methods can be used here.

5.3. APPENDIX - DOMESTIC ABUSE

Over 400 000 incidents of "domestic abuse" ²² were reported to the authorities in the UK in one year in the mid-2000s (BMA Board of Science 2007).

In the British Crime Survey (BCS) 2005/6 (Walker et al 2006) (a victim survey), 363 000 incidents were reported by respondents. Only 42% of them were reported to the police. So official figures are probably under-estimates, particularly as some individuals do not view domestic abuse as a crime (BMA Board of Science 2007).

The abuse can be both active (as in deliberate acts against the victim) and passive (including neglect and lack of concern) (BMA Board of Science 2007).

Domestic abuse has a high rate of repeat victimisation, compared to other violent crime, with over 40% of victims being victimised more than once in the BCS 2005/6 (BMA Board of Science 2007).

BMA Board of Science (2007) listed the barriers to accurate measures of domestic abuse:

- Hidden (and stigmatised) nature of the problem.
- Abuse not viewed as a crime by some perpetrators and victims.
- Domestic violence not distinguished from violence generally in some official statistics.

²² "Domestic abuse" is a wider term than domestic violence embracing violence, and psychological abuse with no physical force (BMA Board of Science 2007).

- Problems of defining "domestic abuse".
- Victims fearful of negative attitudes of authorities (eg: lack of sympathy; viewed as trivial).

5.4. REFERENCES

BMA Board of Science (2007) Domestic Abuse London: British Medical Association

Campbell, J.C (1995) Prediction of homicide of and by battered women. In Campbell, J.C (ed) Assessing the Risk of Dangerousness: Potential for Further Violence of Sexual Offenders, Batterers, and Child Abusers Newbury Park, CA: Sage

Campbell, J.C et al (2003) Risk factors for femicide in abusive relationships: Results from a multi-site case control study American Journal of Public Health 93, 7, 1089-1097

Coakes, S.J & Steed, L.G (2003) SPSS: Analysis Without Anguish (Version 11.0 for Windows) Milton, Australia: John Wiley

Petrie, A (1987) Lecture Notes on Medical Statistics (2nd ed) Oxford: Blackwell Science

Walker, A et al (2006) British Crime Survey 2005/6 London: Home Office