Eyewitness Identification Under Stress in the London Dungeon

Tin Valentine
Jan Mesout

Goldsmiths, University of London.

Corresponding author:

Professor Tim Valentine
Department of Psychology,
Goldsmiths, University of London
New Cross
London SE14 6NW
England

Email: t.valentine@gold.ac.uk

Abstract

Eyewitness experiments do not induce the stress experienced by victims of crime. It is important to understand the effect of stress if results of laboratory studies are to be generalised to victims and witnesses of real crimes, but previous research has shown a mixed picture. The ability of visitors to the London Dungeon to describe and identify somebody encountered in the Horror Labyrinth was investigated, as a function of their state anxiety. To validate the measure of state anxiety, participants wore a wireless heart rate monitor whilst in the Labyrinth. High state anxiety was associated with a higher heart rate. Subsequently visitors completed measures of their state anxiety experienced whilst in the Labyrinth and a measure of trait anxiety. High state anxiety was associated with reporting fewer correct descriptors of the target person, more incorrect details and making fewer correct identifications from a lineup. There was no effect of trait anxiety on eyewitness memory.

Introduction

It is essential for a secure, just society that the criminal justice system is both fair and effective. Notwithstanding procedural safeguards designed to provide a fair trial, evidence shows that wrongful convictions are not uncommon both in the UK and in the USA (Naughton, 2003; Scheck, Neufield & Dwyer, 2000). Mistaken eyewitness identification has been shown to be a common cause of miscarriages of justice in both jurisdictions (Devlin, 1976; Scheck et al, 2000). At the time of writing 215 people wrongfully convicted in the USA have been exonerated by new analysis of DNA from crime scene samples. Mistaken eyewitness identification was a cause of the miscarriage of justice in 75% of these cases (Innocence Project, n.d.).

Evaluation of current procedures and recommendations for best practice in eyewitness identification is almost entirely derived form studies of staged incidents either seen live or on video by participant 'witnesses'. One of the most significant differences between the experience of participant witnesses and the victims of real crimes is likely to be the fear or stress evoked by the crime. For ethical reasons experimental participants cannot be subjected to the same stress that a victim of crime may experience. Nevertheless, it is essential to understand what effect such stress may have on memory in order to understand the processes that may affect the testimony of victims or witnesses of crime.

The literature on the effect of arousal on memory has shown mixed results (e.g. Deffenbacher, 1983). Some studies found that arousal enhances memory while others show a null effect or that arousal impaired memory. Initially the experimental approach was to consider that all stimulation increased arousal along a unidimensional continuum. Effects were interpreted within the context of a U shaped function between performance and arousal - the Yerkes-Dodson law (1908). Increases in arousal associated with the ascending arm were associated with improved memory. Increases in arousal associated with the descending arm were associated with impaired memory. Subsequently an influential review (Christianson, 1992) of the effect of "negative emotionality" on memory rejected this approach, and concluded that memory for central details within the focus of attention is enhanced by increased emotionality but memory for peripheral detail is impaired. Thus, presumed increases in negative affect were assumed to induce a narrowing of the focus of attention.

Results of recent field studies have shown a mixed picture of the effect of stress on eyewitness memory. An often-cited study of eyewitness memory of a real crime suggested that people who experienced more stress recalled more detail of the event (Yuille & Cutshall, 1986). However, the effect of stress in this circumstance was confounded with the degree of involvement in the incident and therefore proximity to details. Police trainees who took part in a stressful simulated stop and search had better eyewitness recall than did trainees who took part in a low stress simulation (Yuille, Davies, Gibling, Marxsen & Porter, 1994). Hulse and Memon (2006) examined recall of police firearms officers who took part in a virtual reality simulation. There were few differences in recall between officers who took part in a 'shoot' scenario in which they fired their weapon compared to officers in a 'no shoot' scenario who did not open fire. Stanny and Johnson (2000) combined a similar virtual reality simulation with a within-participant manipulation of 'shoot' and 'no shoot' scenarios. They found that police and civilian witnesses, who showed increased physiological activation during the violent 'shoot' scenarios, recalled fewer correct details in comparison to their recall of a less violent 'no shoot' scenarios. Ihlebæk, Løve, Eilertson and Magnussen (2003) found that witnesses to a realistic, live, staged bank robbery recalled fewer details than did witnesses who watched the robbery on a video. Peters (1988) found that participants who received an inoculation showed impaired eyewitness identification and recollection of the appearance of the nurse who gave the injection, in comparison to a researcher with whom they interacted for a similar length of time a few minutes after the receiving the injection. The effect was stronger for participants who showed greater physiological activation whilst waiting for their inoculation. Morgan, Hazlett, Doran, Garrett, Hoyt, Thomas, Raranoski and Southwick (2004) examined the ability of soldiers to recall a person present at an interrogation. The soldiers had been detained for 12 hours in a mock prisoner of war camp. Each soldier then underwent, in counterbalanced design, either a high stress interrogation involving physical confrontation followed by a low stress interrogation

without physical confrontation, or the two interrogations in the opposite order. The soldiers were much less accurate in recognizing their interrogator for the high stress interrogation than they were for the interrogator in the low stress condition.

Deffenbacher, Borstein, Penrod and McGorty (2004) argue that in order to understand the effect of stress on eyewitness memory it is necessary to distinguish an activation mode of attention control, which induces increased heart rate, blood pressure and muscle tone, and is likely to be induced in victims of crime, from an arousal mode of attention control, which induces decreased heart rate, blood pressure and muscle tone. They argued that the methodology of the studies reviewed by Christianson (1992) had most likely elicited the arousal mode. Drawing upon a 'catastrophe model', Deffenbacher (1994) argued that at high levels of cognitive anxiety (worry), increasing levels of somatic anxiety (conscious perception of physiological activation) will result in gradual improvement in memory until at some point there is a discontinuous drop in memory performance. Deffenbacher et al. (2004) reported a meta-analytic review of studies in which stress had been successfully manipulated, demonstrated by measures taken as soon as possible after encoding the target person. They found that heightened stress had a moderate negative effect on identification of a target person and on recall of crime-related details. The effect of stress on the rate of correct identification was restricted to target present lineups, there was no effect of stress on the rate of correct rejection of target absent lineups.

The aim of the experiment reported here was to test the prediction derived from the catastrophe model that in a situation which induces cognitive anxiety, high physiological arousal can cause a marked impairment in eyewitness recall and identification. The experiment was designed to test eyewitness memory in a situation that posed some personal threat, but occurs during the normal course of everyday life, without the participants being aware that eyewitness memory would subsequently be tested. First, a subjective self-report measure of state anxiety was validated by use of a measure of physiological activation (heart rate) to establish that the experience was likely to have induced an activation mode of attention control. Recall of the appearance of an actor and identification of him or her from a 'culprit' present lineup served as the dependant measures. Practical limitations of access to the London Dungeon resulted in insufficient participants being available to include both culprit present and culprit absent lineups in the design. Therefore, the experiment was restricted to evaluation of culprit present lineups. After all, previous research has shown an effect of stress on eyewitness identification on target present but not on mistaken identification from target absent lineups (Deffenbacher, et al., 2004). Individual differences in the response to a stressful situation were exploited in a correlational design, rather than attempting to compare a stressful situation with a control condition. A correlational design avoids the problems inherent in attempting to equate the memorability of two different experimental conditions. The experiment was in two parts. First, a subjective self-report measure of state anxiety (Spielberger, 1983) was validated by testing its association with change in heart rate. Second, the eyewitness study was conducted using self-report measures, but without monitoring heart rate, so that participants could be left unaware that their memory would be evaluated.

Method

Participants

Validation of State Anxiety Inventory: Twenty employees of a London department store were recruited. Data from 2 participants was not analysed because

their baseline heart rate was an outlier in the sample. The participants who contributed data for analysis had a mean age of 26.1 years (S.D. = 8.34, range 18 – 48 years), 10 were female, 8 were male.

Eyewitness study: Visitors to the London Dungeon were offered a reduction in the admission price to complete some questionnaires after their visit. Fifty-six volunteers provided complete data that were analysed. They had a mean age of 31.0 years (S.D. = 9.15, range 18 – 54 years), 29 were female, 27 were male.

Materials

The Horror Labyrinth is the first exhibit in the tour of the London Dungeon. It is a maze of floor to ceiling mirrored walls set amongst Gothic vaults. Visitors are admitted in groups of around thirty people. Designed to disorientate visitors, it is dark, crowded and there is a sound track playing the rhythm of a heart beat and various scary noises and screams. Walking though the maze, visitors often find an apparent path through the arches suddenly blocked by a mirror. There are a number of 'scares' in the Labyrinth. A screaming skeleton is triggered by an infra-red beam when approached. An elderly woman sitting in a rocking chair appears behind a mirror. There was an actor in the Labyrinth, dressed in a dark robe and wearing theatrical make up to create a very pale facial skin colour with wounds or scars. This actor, hereafter referred to as the 'scary person', would step out in front of the participant, and then block their path to prevent them passing.

Heart rate was measured using a Polax Accurex Plus wireless heart rate monitor. State and Trait anxiety was measured using the Spielberger (1983) State-Trait Anxiety Inventory (Form Y). The state anxiety inventory (SAI) consists of 20 items. Each statement was altered to be worded in the past tense. For example "I was tense", "I felt frightened". Each statement was rated on a four point scale (not at all, somewhat, moderately so, very much so). A score between 1 and 4 was assigned for each item. The trait anxiety inventory (TAI) consisted of 20 statements relating to how one usually feels, such as "I feel nervous and restless", "I have disturbing thoughts", "I feel inadequate". Each statement was rated on a four point scale (almost never, sometimes, often, almost always) and assigned a score between 1 and 4. A questionnaire was designed to record free recall and cued recall of the scary person encountered in the Labyrinth. Photographs of 50 actors from the London Dungeon were used to compile a nine person photograph lineup for the actor working in the Labyrinth.[1] The foils most similar in appearance, matched on sex, approximate age and ethnic origin to the scary person, were selected from the 50 photographs available. The foils wore similar theatrical make-up including wounds or scars and similar costume to the scary person. All photographs were full-face views, taken under similar conditions with the same background. The actor playing the scary person differed from day to day. A custom lineup was selected on the day for each actor. Most of the 50 actors played the role of the 'scary person' on different occasions.

Design

Validation of SAI: An average baseline heart rate was measured with the participants walking slowly. Immediately afterwards their average heart rate was recorded whilst the participant was in the Labyrinth with other visitors, walking slowly. Having completed their visit to the London Dungeon, the participants reported, using the SAI, how they had felt whilst in the Labyrinth. The difference in

[1] A nine person lineup was used because identification in the UK, which is based on a video procedure, requires a minimum of 8 foils and a suspect.

the two average heart rates was correlated with the state anxiety score to measure the association of the SAI score with change in heart rate.

Eyewitness study: A multiple regression design was used. Three separate analyses were conducted on the dependent variable, the number of descriptors recalled correctly, incorrectly reported descriptors, and the outcome of the identification procedure. The independent variables were sex, state anxiety score (whilst in the Labyrinth) and trait anxiety score. Accuracy on the eyewitness identification from a photograph lineup was also analysed as a function of SAI, by using the median SAI score to divide participants into a high state anxiety group and a low anxiety group.

Procedure.

Validation of SAI: Informed consent was obtained before taking part in the study. The wireless heart rate monitor was strapped around the participant's chest. Their baseline heart rate was recorded whilst walking slowly for 7 minutes. The participant then entered the labyrinth with a group of other visitors. The participant was pointed out to the actor to ensure that she or he encountered the scary person whilst walking slowly around the labyrinth for approximately 7 minutes. Each participant then completed the tour of the London Dungeon visiting the other exhibits in a fixed order. About 45 minutes after leaving the labyrinth the participant completed the state anxiety questionnaire. They were instructed to read each statement and to circle the response which best described "how you felt when you were in the labyrinth".

Eyewitness study: Participants in the main part of the experiment had agreed to fill in some questionnaires about their visit but had not given informed consent to participate in the experiment. Each had visited the Labyrinth in exactly the same manner as described above for the participants in the validation study. Each participant was pointed out to the actor in the Labyrinth to ensure that she or he encountered the scary person whilst walking slowly around the labyrinth for approximately 7 minutes. About 45 minutes later, after they had completed their tour visiting the other exhibits in the same fixed order, the purpose of the experiment was explained and informed consent was obtained. It was explained that the participant was free to withdraw at this or any other point without having to give a reason. The participants who gave consent completed the SAI, and were instructed to circle the response which best described "how you felt when you were in the labyrinth". They then completed the TAI, for which they were instructed to read each statement and to circle the response which best described "how you generally feel". Then they completed a questionnaire on their memory for the 'scary person'. First they were asked to provide a written free recall of the description of the scary person. Then they were prompted to provide a cued recall (e.g. sex, age, height, hair colour, clothing etc.). In both recall tasks participants were instructed to include only those details they could actually remember and try not to guess any details they were unsure of. Finally participants were shown a nine person photograph lineup. All pictures were shown simultaneously in two rows (one with 4 pictures and one with 5 pictures). The picture of the 'scary person' was placed in a randomly selected position for each lineup. Unbiased instructions were used. Each participant was instructed that "the person you saw in the labyrinth may or may not be in the lineup", and that "if you cannot make an identification you should say so". After making their response, participants rated their confidence in their decision on a scale of 0 – 100% confidence.

Results.

Validation of SAI: Data from two participants were excluded because exploratory analysis showed that their baseline heart rate was an outlier in the distribution of the sample (> 98 beats per minute [bpm]). The mean baseline heart rate of the remaining 18 participants was 74.7 bpm (s.d.= 6.3). Their mean heart rate while in the labyrinth was 86.9 bpm (s.d = 9.3). Mean heart rate was reliably higher in the labyrinth, $t(17)=6.4$, $p<.001$, $d=1.53$. Their mean state anxiety score was 43.2 (s.d. = 2.4). Pearson's correlation coefficient between the change in heart rate (Labyrinth mean heart rate – baseline mean heart rate) and state anxiety score showed a reliable association, $r=.76$, $p<.001$, $r^2=.58$. Change in heart rate accounted for 58% of the variance in state anxiety score.

Eyewitness study: Complete data were obtained from 56 participants. The mean state anxiety score was 49.0. State anxiety was reliably higher for females than for males (52.8 vs. 45.3 respectively), $t(54) =2.84$, $p<.01$, $d= 0.75$. The mean trait anxiety score was 36.8. There was no difference in trait anxiety between males and females (36.3 vs. 37.3).

Descriptors provided in the free and cued recall tasks were divided into single units of information (noun and adjective phrases) and combined into a single score. Two judges rated each descriptor as correct or incorrect compared to a description noted of the appearance of each actor, including their clothing and make-up on the day. Each actor supplied self-report data on their age, height, weight, eye colour and hair colour. A photograph of the actor was used to evaluate descriptions of facial features. Estimates were judged correct if height was +/- 2 inches, age was +/- 2 years, weight was +/-5 pounds. Statements with qualifiers e.g. 'might' etc. were scored as incorrect following the procedure adopted by Yuille and Cutshall (1986). Two scorers worked independently. They agreed in their classification of items as correct or incorrect for 94% of items. Disagreements were resolved by discussion.

The simple correlations between the number of correct descriptors given by each participant and their sex, state anxiety score and trait anxiety score are shown in Table 1. Participants who reported lower state anxiety recalled more correct descriptors, $r=-.41$, $p=.001$, $r^2= .17$. Females reported higher state anxiety than did males, $r=.36$, $p <.005$, $r^2= .13$.

Variables	1 Ncr.	3 sex	4 state	5 trait.	*B*	β	Partial
1. (DV) Ncr.	1.00	-.08	-.41**	.11		.	.
3. Sex		1.00	.36*	.08	.42	.07	
4.State			1.00	.09	-.13	-.45**	-.42**
5. Trait				1.00	.06	.14	
						Constant = 9.6	
Means	6.25	1.48	49.0	36.8			
Sd.	3.01	0.50	10.5	6.54			
							R^2 = .19
	Adjusted R^2 = .15						*R* = .44**
							S.E. = 2.8

Ncr = Number of correct descriptors * p<.005 **=p<.001

B = Unstandardize Regression Coefficients; β = Standardised Regression Coefficients

Table 1: Multiple regression of the number of correct descriptors of the target person (Ncr).

The number of correct descriptors recalled was entered as the dependent variable in multiple regression. The predictor variables entered were sex, state and trait anxiety. The results are shown in Table 1. People who reported higher state anxiety recalled fewer correct details, β=-.45, t=-3.34, p<.001, partial correlation = -.42, p<.001. No other effects were significant. Examination of the colinearity statistics indicated that there was not a problem with colinearity in the model.

A second multiple regression model was constructed to predict the number of incorrect descriptions reported. The simple correlation with sex, state anxiety score and trait anxiety score are shown in Table 2. There is a reliable association between the number of incorrect descriptors and state anxiety, r=.53, p<.001, r^2= .28. Participants who reported higher state anxiety reported more incorrect descriptors.

The number of incorrect descriptors reported was entered as the dependent variable. The predictor variables entered were sex, state and trait anxiety. The results are shown in Table 2. People who reported higher state anxiety reported more incorrect details, β=-.51, t=4.14, p<.001, partial correlation = .47, p <.001. No other effects were significant. Examination of the colinearity statistics indicated that there was not a problem with colinearity in the model.

Variables	1 Nic.	3 sex	4 state	5 trait.	*B*	β	Partial
1. (DV) Nic	1.00	.27	.53**	-.12		.	.
3. Sex		1.00	.36*	.08	.47	.10	
4.State			1.00	.09	.09	.51**	-.47**
5. Trait				1.00	-.05	-.17	
						Constant = 9.6	
Means	2.80	1.48	49.0	36.8			
Sd.	1.93	0.50	10.5	6.54			
							R^2 = .32
	Adjusted R^2 = .28						*R* = .56**
							S.E. = 1.6

Nic = Number of incorrect descriptors * p<.005 **=p<.001
B = Unstandardize Regression Coefficients; β = Standardised Regression Coefficients.

Table 2: Multiple regression of the number of incorrect descriptors of the target person (Nic).

The outcome of attempted identification from a 9 person photograph array, coded as accurate or inaccurate, was analysed by logistic regression. Accuracy of identification was the dependent variable, sex, SAI and TAI entered at step 1 as independent variables (–2 log likelihood of the model = 56.4, Nagalkerke R square = .42). There was a reliable effect of state anxiety, Wald(1)= 9.40, p<.005. The effect of sex was just short of a conventional level of significance, Wald(1)= 3.56, p=.059.

To explore the effect of state anxiety further a median split of state anxiety scores (51.5) was used to divide participants into high and low state anxiety groups (see table 3). Expected frequencies for the no identification category were less than 5. Therefore, the outcome of identification was divided into accurate and inaccurate (no identification or a foil identification) for analysis. Participants who reported high state anxiety were less likely to correctly identify the person in the labyrinth, $\chi^2(1)=12.95$, p<.001, Cramer's V = .52. In view of the association between state anxiety and sex, identification accuracy was also analysed as a function of sex (see Table 4). Males

made more correct identifications than did females, $\chi^2(1)=8.81$, $p<.005$, Cramer's V = .40.

	Low state anxiety	High state anxiety	Total
Culprit ID	21	5	26
Foil ID	6	15	21
No ID	1	8	9
Total	28	28	56

Table 3: The outcome of identification attempts from a target present nine-person photograph lineup as a function of state anxiety.

	Male	Female	Total
Culprit ID	19	7	26
Foil ID	5	16	21
No ID	5	4	9
Total	29	27	56

Table 4: The outcome of identification attempts from a target present nine-person photograph lineup as a function of sex.

For participants who made an identification from the lineup, accuracy of identification was reliably associated with confidence rated prior to the identification attempt, $r(47) =.39$, $p<.001$, $r^2= .15$, and to confidence rated after making the attempted identification $r(47) =.58$, $p<.001$, $r^2= .34$. The accuracy-confidence relationship could not be calculated for non-choosers because only target-present lineups were used, and therefore all non-choosers incorrectly rejected the lineup.

Discussion

The experience of the Horror Labyrinth at the London Dungeon was successful in inducing physiological arousal, evidenced by a reliable increase in heart rate in the validation group. Furthermore, the state anxiety scores for participants in the eyewitness study were elevated (mean = 49.0) compared to norms for working adults aged 19 – 39 (36.5 for males, 36.2 for females; Spielberger, 1983). The trait anxiety scores (mean = 36.8) were very similar to the Spielberger (1983) norms (35.5 for males and 36.1 for females). These data show that trait anxiety of participants was average but the Horror Labyrinth induced both physiological activation (i.e. elevated heart rate) and the subjective experience of anxiety.

There was a reliable association between sex and state anxiety, with females reporting higher state anxiety in the Labyrinth than did males. This effect is consistent with previous data which show females reported higher state anxiety under stressful conditions (exam, film of injuries) than did males (Spielberger, 1983).

There was a strong negative association between state anxiety and the ability to correctly report the appearance of a person encountered under stressful conditions. Witnesses who experienced higher anxiety reported fewer correct descriptors than witnesses who experience less anxiety. This effect is consistent with the catastrophe model (Deffenbacher, 1994); at high levels of physiological activation and cognitive anxiety eyewitness memory is impaired. There was also a strong association between state anxiety and the number of reported descriptors judged to be incorrect. The scoring scheme used coded descriptors recalled with qualifiers such as 'might have', 'I think' etc. as incorrect. Therefore this association may reflect a tendency for high

state anxiety participants to report details with more qualifiers than low state anxiety participants, rather than reporting details that are necessarily mistaken.

Eyewitness identification was dramatically impaired by high state anxiety. Only 17% of eyewitness who scored above the median on the state anxiety scale (52 or above) correctly identified the person they saw from a nine-person culprit-present photograph lineup. In contrast 75% of eyewitnesses who scored below the median correctly identified the 'culprit'. The effect of state anxiety found in the present study is consistent with the results of Morgan et al. (2004) who found a similarly catastrophic impairment of eyewitness identification after participants underwent a high stress interrogation. A catastrophic impairment of eyewitness identification in situations of high stress has also been reported by Peters (1988), amongst participants who received an injection, and by Bothwell, Brigham and Piggott (1987). Bothwell et al. found that participants who scored highly on a neuroticism scale showed catastrophic impairment of eyewitness identification under conditions of high stress induced by the presence of a syringe lying on a table at which the participant was seated. The participants in these three previous studies and the present experiment were unaware that they would be required to make an eyewitness identification from a situation that posed some personal threat.

The strength of the accuracy-confidence relationship for identification is similar to the effect size reported previously in the literature when the calculation is restricted to choosers (Sporer, Penrod, Read & Cutler, 1995). Post-identification confidence showed a stronger association with accuracy than did pre-identification confidence. This result is also consistent with previous research (Cutler & Penrod, 1989).

The association between state anxiety and sex meant that eyewitness identification under the stressful conditions in the Horror Labyrinth was much less accurate when made by female witnesses than by male witnesses. Sex differences were still present but less reliable when the effect of state anxiety is accounted for separately. Sex differences in face recognition are generally small, but favour females: women recognise faces more accurately than do males (Shapiro and Penrod, 1986). An implication of the results and the association of state anxiety under stress with sex is that eyewitness identification may be particularly vulnerable to the effect of stress in female witnesses or victims of crime.

The study has demonstrated that a physiological measure of arousal (increased heart rate) is closely associated with a subjective report of anxiety (SAI). Furthermore the experience of stress in a naturalistic context had a similar, catastrophic effect on the accuracy of eyewitness memory to that observed in more artificial laboratory studies that successfully induce an activation mode of attentional control (Deffenbacher et al., 2004). These data suggest that the accuracy of eyewitness identification by some victims of crime has been over-estimated in data from laboratory studies that do not induce an activational mode of attentional control. This conclusion suggests that the potential for error associated with eyewitness identification evidence is even greater than laboratory studies have suggested.

The catastrophic failure of the ability to describe and identify a person encountered under high state anxiety in the Horror Labyrinths provides further strong support for the catastrophe model of memory performance under anxiety (Deffenbacher, 1994) and the analysis of the literature described by Deffenbacher et al. (2004). The strategy of recruiting people from a tourist attraction that induces high state anxiety provides a useful methodology for study of memory under stress in a naturalistic setting. In common with victims or witnesses of crime, our participants

had no prior warning that they would be required to identify the stranger they had met and the encounter occurred under a situation that elicited heightened physiological activation and cognitive anxiety. This approach takes advantage of individual differences to study the effect of stress rather than comparing a stressful situation with a control condition. This avoids the problem of equating the memorability of the stressful and control conditions. Future research could explore further the limits of the negative impact of stress on identification. Suitable situations to study include memory for people encountered at stressful fairground rides (e.g. the person loading the cars on the big dipper) or extreme sports (e.g. an instructor at a bungee jump).

References

Bothwell, R. K., Brigham, J. C. & Pigott, M. A. (1987). An exploratory study of personality differences in eyewitness memory. Journal of Social Behavior and Personality, 2, 335-343.

Christianson, S-Å. (1992). Emotional stress and eyewitness memory: A critical review. Psychological Bulletin, 112, 284-309.

Cutler, B. L. & Penrod, S. D. (1989). Forensically-relevant moderators of the relationship between eyewitness identification accuracy and confidence. Journal of Applied Psychology, 74, 650-652.

Deffenbacher, K. A. (1983). The influence of arousal on reliability of testimony. In S. M. A. Lloyd-Bostock & B. R. Clifford (Eds.). Evaluating witness evidence. Chichester: Wiley. (pp. 235-251).

Deffenbacher, K. A. (1994). Effects of arousal on everyday memory. Human Performance, 7, 141-161.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D. & McGorty, K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. Law and Human Behavior, 28, 687-706.

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

Hulse, L. & Memon, A. (2006). Fatal Impact? The Effects of Emotional Arousal and Weapon Presence on Police Officers' Memories for a Simulated Crime Legal & Criminological Psychology, 11, 313-325

Ihlebæk, C., Løve, T, Eilertson, D. E., Magnussen, S. (2003). Memory for a staged criminal event witnessed live and on video. Memory, 11, 319-327

Innocence Project (n.d). Retrieved March 27, 2008, from Innocence Project website: http://www.innocenceproject.org/understand/Eyewitness-Misidentification.php

Morgan, C. A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baranoski, M. & Southwick, S. M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. International Journal of Law and Psychiatry. 27, 265-279.

Naughton, M. (2003). How big is the 'iceberg'? – a zemiological approach to quantifying miscarriages of justice. Radical Statistics, 81, 5 –17.

Peters, D. P. (1988). Eyewitness memory and arousal in a natural settting. In: Gruneberg, M. M., Morris, P. E. and Sykes, R. N. (eds.) Practical aspects of memory: Current research and issues. *Vol. 1*. Chichester: John Wiley & Sons. (pp. 89-94).

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

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

Spielberger, C. D. (1983). State-Trait Anxiety Inventory (Form Y). PaloAlto: Mind Garden.

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

Stanny, C. J. and Johnson, T. C. (2000). Effects of stress induced by a simulated shooting on recall by police and citizen witnesses. American Journal of Psychology, 113, 359-386.

Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. Journal of Comparative Neurology and Psychology, 18, 459-482.

Yuille, J. C. & J. L. Cutshall (1986). A case study of eyewitness memory of a crime. Journal of Applied Psychology, 71, 291-301.

Yuille, J. C., Davies, G., Gibling, F., Marxsen, D. & Porter, S. (1994). Eyewitness memory of police trainees for realistic role plays. Journal of Applied Psychology, 79, 931-936.

Acknowledgement.

We thank Mark Oakley of Merlin Entertainments and the staff of the London Dungeon for their assistance in making arrangements for data collection in the Horror Labyrinth, and Ken Deffenbacher for helpful comments on an earlier draft of the manuscript.