"TRUTH, THE WHOLE TRUTH AND NOTHING BUT THE TRUTH?" FACTORS INFLUENCING EYEWITNESS TESTIMONY

Justyna Stachowiak
London University, London

Summary. Several factors considered to be reliable indicators of eyewitness accuracy were examined. It was hypothesised that personality traits and position on the field dependence/independence dimension would be two prerequisites that alter suggestibility of the eyewitness. Additionally, the relationship between suggestibility to leading questions and reported confidence of the answers was examined. Thirty adults viewed a video of a staged shop robbery not knowing about a memory recall task. Participants completed Eysenck's Personality Questionnaire and Embedded Figures Test. A new questionnaire was constructed to measure suggestibility. It consisted of 13 normal and five leading questions. Participants' confidence in their answers was assessed on a separate scale. Univariate ANOVA was performed on individual suggestibility scores which were also correlated with confidence rating, overall EFT score and scores on three personality scales (introversion/extraversion and neuroticism). The results suggest that with regards to leading questions introvert participants with field dependent style have more difficulty providing the correct answer than extrovert, field independent participants. Correlations obtained for confidence suggest that it is a trait independent of suggestibility. Reported levels of confidence were similar for subjects who provided correct and incorrect answers to leading questions. The finding re-emphasised the importance of using correct methods to elicit eyewitness evidence. The results are discussed in relation to applying knowledge about factors influencing eyewitness evidence in the forensic settings.

Introduction

The jury in many trials heavily relies on eyewitness testimony. It is therefore most important that witnesses give complete and accurate evidence in order to maximise the probability of a guilty defendant being convicted or an innocent defendant being acquitted. A number of research projects have found that the way in which the witness is interviewed by police has a great impact on accuracy of the testimony (Loftus, 1975). However, Kebbel and Johnson (2000) highlight that the way in which a witness is questioned in the court room has been researched on a much smaller scale. It is the

---

Adres do korespondencji: just_tynka@yahoo.co.uk

task of the police investigators to elicit the full version of the events. Counsel, whether prosecution or defence, only needs to obtain the part or version of the evidence that will serve his or her case best (Kebbel and Glichrist, 2004). Knowing that the prosecution needs to prove guilt beyond reasonable doubt and that the defence has to merely prove that the doubt exists, the jury often have a very difficult task.

Indeed, contradictory results have been found with regards to the type of evidence presented to the jury (physical *vs.* eyewitness). A large body of research shows that jurors heavily depend on eyewitness accounts of events (e.g. Cutler, Penrod, Dexter, 1990; Wells, Lindsay, Ferguson, 1979). However, Wells's review (1998) estimates that prior to DNA testing approximately 75% of convicts were victims of mistaken eyewitness identification. The most common reason for false conviction is incorrect witness evidence. Interestingly, Skolnick and Shaw (2001) showed that mock-jurors would not overemphasise witness evidence if physical evidence was strong. This would suggest that investigators should turn their attention to physical evidence which is tangible and much more scientific. Unfortunately, in many cases eyewitness evidence is the only available source of information.

Psychologists and other specialists have been investigating a large number of factors that might influence eyewitness accuracy. These are usually classified as characteristics of the witness, the target, the situation and testimony, which include post-event information and techniques (e.g. Narby, 1996). In their review Narby et al. identified 34 different factors that may have an impact on the evidence. Wells (1978) categorised those factors from the justice system's point of view. He defines estimator variables as factors over which the justice system has little or no control, i.e. personality traits or level of stress experienced while witnessing crime. Factors like these cannot be controlled outside the laboratory and in a real-life situation they can only be estimated post hoc. The other type – system variables – can be easily manipulated by the criminal justice system and are less likely to be biased by the witness. These include selection of line-up members and questioning techniques (Narby, 1996).

The focus of the present research was factors from the first category – estimator variables – and included personality traits (namely extraversion/ introversion and neuroticism), cognitive style (field dependence/independence), suggestibility and confidence reported post hoc. The last two factors would be classified by Wells (1978) as system variables because they can be induced by confusing questions which is decided and controlled by lawyers. However, suggestibility and confidence are characteristic of the witness and were thus labelled as estimator variables. The study aimed to investigate relationships between these factors and their overall impact on eyewitness testimony.

A major factor that can greatly influence original testimony is suggestibility of the witness. It can be defined as the "tendency to provide the answer believed to be required by the questioner" (Kebbel et al., 2004) and is often induced by leading questions (e.g. The car was blue, wasn't it?). This type of question is commonly used by barristers because a strong suggestion of the expected answer has a confusing effect on a witness.

In terms of experimental testing of suggestibility there are two main traditions. Gudjonsson (1984b) developed his model of suggestibility based on the "individual differences approach", which focused on coping strategies during the interview. The Gudjonsson Suggestibility Scale (GSS) was designed to measure confidence about correct answers, trust of the interviewer, and expectation of success. In contrast, the "experimental approach" is based on manipulating the situation in order to understand the cognitive processes associated with suggestibility. The Classic design (e.g. Loftus, Zanni, 1975) involves participants being shown a stimulus followed by misleading information about it and testing memory of the original event. Critics of laboratory research argue that suggestibility is a multidimensional concept that is very difficult to measure and may easily be overestimated. Laboratory conditions are able to simulate memory tasks but lack factors such as stress, distraction or fear and therefore are not representative of a real-life situation faced by the witness (Ihelbaek, Love, Eilersten, Magnussen, 2003). According to Lane (in press) if attention is divided during encoding, the witness is more suggestible to post-event information. In laboratory conditions subjects usually pay their full attention to the stimuli.

Pezdek and Greene (1993) suggested that level of suggestibility depends on similarity of the retrieval environment to the original encoding environment. In the large majority of eyewitness research items in the recall session have been presented in a random order which resulted in high suggestibility (e.g. Bekerian, Bowers, 1983). Despite those findings, the present study created a suggestibility scale with items in an order that matched the order of events seen during input. The underlying aim of this study was to match as closely as possible the actual conditions in which a witness would give evidence and this was partly achieved by retaining the chronological order of the questions. It was expected that this decision would result in lower levels of suggestibility when compared with similar studies but which employed random ordering.

Overall witnesses' resistance to leading questions remains a very reliable measure of suggestibility and has been found to correlate with other factors. Stern, as early as 1926 (see: Endres, 1997) reviewed several forensic cases and developed an idea that suggestibility depends both on characteristics of the witness and of the interview situation. Gudjonsson (1983) using his new suggestibility questionnaire, Wechsler Adult Intelligence Scale and the Eysenck Personality Questionnaire, found suggestibility significantly negatively related to intelligence, and positively related to social desirability and neuroticism.

The relationship between individual differences in personality and eyewitness performance has not been explored to a great extent. The reason for that might be the complexity of personality. Also, it cannot be manipulated prior to witnessing crime in order to alter accuracy of evidence (Pozzulo, Coplan, Wilson, 2005).

Extraversion/introversion and neuroticism/stability, Eysenck's original dimensions, are one of the most commonly used ways to classify and describe personality. Research suggests that "neuroticism (N factor) is closely related to the inherited degree of lability of the autonomic nervous system, while extraversion (E factor) is closely related to the degree of excitation and inhibition of the central nervous system"

(Eysenck, Eysenck, 1991). These biological explanations can be loosely translated into descriptions of "typical" extravert and introvert. The former would be sociable, impulsive, easygoing and optimistic; the latter would be quiet, introspective, planning ahead and having high ethical standards. A person scoring high on the N scale could be described as anxious, moody and worrying whereas a low scorer would be calm, even-tempered and controlled. In terms of eyewitness accuracy it would be straightforward to assume that an open and impulsive extrovert would make a better witness than an introspective person. However, there are some very interesting data available. In research by Ward and Loftus (1985) no significant difference in accuracy between extraverts and introverts was found although introverts were slightly more susceptible to post-event information. One of the proposed explanations was that introverts have less confidence in their memories and are more willing to rely on information from other sources. However, according to Trouvé and Libkuman (1992) "eyewitness performance was superior for introverts relative to extroverts". On the other hand, in their study of personality characteristics Gudjonsson, Sigurdson, Bragason, Einarsson, and Vladimarsdottir (2004) found that compliance correlated positively with neuroticism and negatively with extraversion. The current study attempted to replicate these findings in relation to suggestibility.

Field dependence-independence (FDI) is another personality characteristic, but related to perceptual and cognitive information processing styles. FDI is a perceptual construct that runs on a continuum, from high field dependence (FD) to high field independence (FI). Individuals described as FD have difficulty in separating information from the organisation of the surrounding field and are less selective in their information intake. FI subjects are better at separating parts of the field from it's context and extract essential information faster (Witkin, Dyk, Faterson, Karp, 1971).

A person's position on this continuum can be measured by two standard tests. The Rod and Frame Test (RFT) requires participants in a dark room to position vertically a tilted luminous rod, irrespective of a tilted frame within which it is placed. The Embedded Figures Test (EFT) requires participants to locate simple, geometric figures that are embedded within more complex ones. Individuals that have few problems with this task can be classified as FI and those who cannot easily separatethe simple form are FD. The difficulty experienced by FD individuals reflects an effect of the surrounding context. This can have important effects in real-life situations where, for example, FD individuals would tend to encode the context of an event, while FI individuals would tend to memorise the essential details (Emmet, Clifford, Gwyer, 2003). FDI scores are very consistent when measured on both RFT and EFT, thus this cognitive style is thought to be a very stable characteristic (Witkin et al., 1971).

An important implication for criminal investigation is that knowledge of a person's cognitive style can be used to support them during recall and so help them perform better. FD subjects benefit from context reinstatement during free recall, because they encode the environmental cues of the event. FI subjects tend to perform better on cued recall (Emmett et al., 2003). Smith and Rothkopf (1984) argue that focus on the environmental context is a reason for higher susceptibility to sug-

gestions of FD witnesses. This report aimed to replicate these findings for free recall without context reinstatement.

Taken as a whole, research implies that suggestibility can be easily induced by the form of questioning and that, during standard investigation, introvert witnesses with a tendency towards field dependence would be more prone to it. However, despite those strong prerequisites and clear connection between them, Kebbel and Johnson (2000) suggest that damage done to the testimony might possibly be reduced by maintaining high levels of confidence. According to Cutler et al. (1990) the apparent confidence of the witness is a very important source of information for the jury. If, due to lawyers' questioning a witness also lacks confidence, the evidence is likely to be treated as not dependable, whether or not the accuracy is impaired. The relationship between confidence and accuracy can be changed easily. This is due to a fact that retrieving memories and reporting confidence are cognitive processes independent of each other (Shaw and McClure, 1996). Thus, accuracy can remain high, while confidence is impaired, or the opposite can happened. There is much evidence for that. Repeated questioning after the event does not affect accuracy but leads to overconfidence, possibly by forming a familiarity effect (Shaw, McClure, 1996). Non-specific feedback from investigators can also have an effect on memories, that is particularly important when dealing with vulnerable or young witnesses (Hafstad, Memon, Logie, 2004). A relationship between confidence and personality measures has also been found. Trouvé et al. (1992) reported higher confidence in extraverts in comparison with introverts and a positive relationship between confidence and accuracy.

Since confidence clearly is a subjective measure, some researchers have suggested that pre-event questioning would be a better indicator of it's levels. It was found that an a priori check of confidence correctly indicates ability to be a good witness, whereas a post hoc measure did not correlate with actual accuracy (Berger, Herringer, 1991). Despite theoretical and experimental advantages, this approach has little use for the criminal justice system because witnesses are judged on their confidence while giving evidence. Also, for obvious reasons, in a forensic situation confidence cannot be measured before an event. To reflect reality, in this study post-event confidence was measured together with suggestibility.

The overall aim of this study was to assess the influence of various factors on the accuracy of witnesses and to investigate the relationships between them.

With regards to suggestibility to leading questions it was expected that witnesses would vary in performance, with some falling for the leading questions and some providing the correct answer. It was hypothesised, that this difference would be related to performance on personality and performance on cognitive tasks. Extrovert subjects with low neuroticism scores would be less prone to misinformation than introvert subjects high on the neuroticism scale. With respect to FDI positioning, FD subjects were expected to have greater suggestibility to leading questions than FI subjects.

A confidence variable was an addition to the main design. The literature did not enable specific predictions to be made about a relationship between suggestibility and confidence for either leading or non-leading questions.

Method

Participants. Thirty participants (19 female, 11 male) took part. Eighteen subjects were recruited among the author's acquaintances, and 12 were undergraduate psychology students. All participated voluntarily and signed a consent form. The mean age was 34.6 years (range 22 – 56 years, $SD = 10.19$).

Materials. A staged video of Blockbuster Video shop robbery was made available by Dr F. Gabbert (Gabbert, Memon, Allan, Wright, 2004). The film depicts an employee standing behind the till. A customer enters the shop, asks a few questions and browses through the videos. Two men enter the shop, disguise their faces and pursue a robbery. A customer is pushed and falls down. The Robbers leave the shop with the money. The film lasts 1 minute, 25 seconds.

A list of eighteen questions concerning the movie was constructed (*Appendix 1*). There were 13 factual questions that could be answered after seeing the film. In order to test suggestibility, five leading questions were added (numbers 1, 5, 13, 15 and 17). They contained strong suggestions that an event that did not take place in the film, actually happened, and asked details about it (e.g. "Did the robbers steal a video from 'action' or 'new releases' section?").

The order in which the questions were asked was not randomized, so as to reflect the reality of giving evidence. During the court hearing a witness is usually asked about the events in a chronological order.

Additionally, a 5-point confidence scale was employed; it ranged from "not confident at all" (1) to "very confident" (5).

Eysenck Personality Scales, EPS (Eysenck, Eysenck, 1991) consisting of 106 questions was used to assess three personality dimensions:
1) introversion vs. extraversion
2) psychoticism vs. impulse control
3) neuroticism vs. emotional stability

Embedded Figures Test, EFT (Witkin et al., 1971), was used to assess cognitive style, field dependence – independence (FDI).

Procedure. Participants were told they would watch a short video of a staged robbery. It was stressed to them that it was a simulated event. They were instructed to pay attention but they were not told they would be tested on it. Afterwards subjects were asked to complete the EPS which also served the purpose of a filler task to create a delay between observing the video and answering questions. That was followed by the suggestibility questionnaire, which asked participants questions about the film. Written instructions were reinforced by vocal instructions and stressed that every question should be answered according to what he/she remembered from the video. The confidence scale was also explained. After completion, the form was checked by the researcher for missing answers or confidence ratings and participants were asked about unclear answers. The EFT was then administered. Subjects were shown the booklet, were given time to read the instructions and practise on two example tasks. The most important instructions were stressed vocally. The EFT consists of three sections and participants were allowed maximum of 5 minutes to complete each section.

A full testing session lasted 35-40 minutes. At the end of it participants were briefed on the purpose of the experiment.

Design. A correlational design was employed with participants being assessed on field dependence-independence, extraversion-introversion, suggestibility and confidence. An additional intention was to create a $2\times2\times2$ *quasi*-experimental design on the basis of FDI, extraversion-introversion and suggestibility.

Results

All 18 items in the suggestibility questionnaire were scored. There were 13 non-leading questions. Five of them had only two possible answers and were straightforward to score (e.g. Did the cashier open the till on his left or right first?). Eight normal questions required a descriptive answer, therefore sometimes there could be several correct outcomes. For example to question no. 11 ("How did the main robber get the employee to hurry up?") correct answer could be: by shouting, by throwing things, by hitting the computer. An answer to a leading question (e.g. "Did the robbers steal a video from 'action' or 'new releases' section?") was only scored as correct if the participant wrote an answer such as: "They didn't steal a video", instead of choosing one of the options suggested in the question. Questions with no answers given were also counted. Therefore six categories were created: leading questions answered correctly and incorrectly, normal questions answered correctly and incorrectly and leading and normal questions not answered. Mean score for every category of questions was calculated for every participant. For every category of questions a relevant mean confidence was also calculated.

The median EFT score of 13 was used to divide participants into two groups. Those whose scores fell below the median were classified as field dependent (FD), whilst those whose scores fell on or above the median were classified as field independent (FI).

A univariate ANOVA was performed to examine the average number of leading questions answered incorrectly. The two between participant conditions were FDI classification (FD *vs.* FI) and EPI classification (extravert *vs.* introvert). The mean scores and standard deviations for the four conditions are presented in Table 1.

Table 1. Number of leading questions answered incorrectly

EFT score	EPQ score	Mean	Std. D.	N
Field dependent	introvert	4.14	1.07	7
	extravert	2.14	1.46	7
	Total	3.14	1.61	14
Field independent	Introvert	3.00	1.79	6
	extravert	2.30	1.77	10
	Total	2.56	1.75	16

The ANOVA indicated that the main effect of EPI classification was significant, $F_{(1.26)} = 5.37$, $p = 0.02$, with extravert participants performing slightly better than introvert ones. The main effect of FDI classification was not significant, $F_{(1.26)} = 0.717$, $p = 0.4$, however field-dependent subjects seemed to fall for leading questions more often than field-independent subjects. The interaction between field dependence/independence and extraversion/introversion was not significant, $F_{(1.26)} = 1.24$, $p = 0.27$. Despite lack of significant interaction, from Figure 1 it can be seen that for field dependent participants, introverts answered fewer leading questions incorrectly than did extraverts.
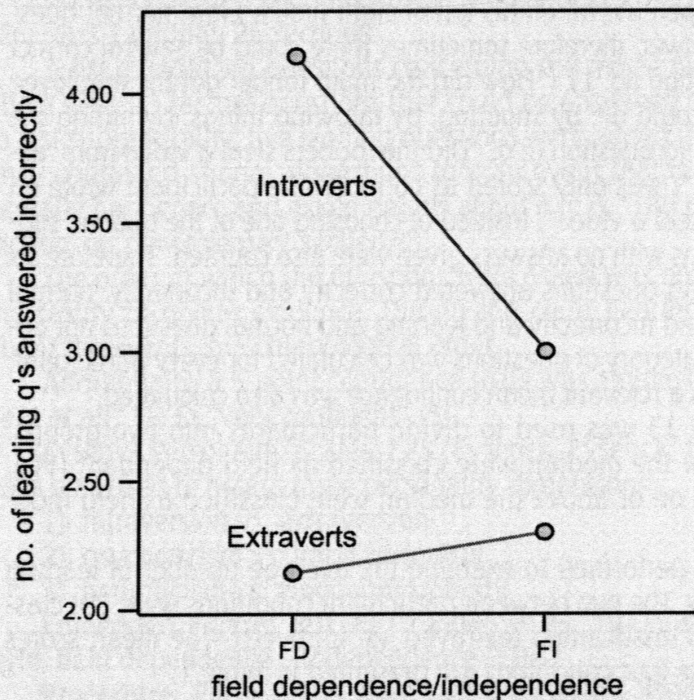
Figure 1. Performance of introverts and extraverts on number of incorrect answers to leading questions and FDI

An additional ANOVA with the same between subject factors was performed to examine the average number of normal questions answered correctly. Neither the effect of FDI classification, $F_{(1.26)} = 0.751$, $p = 0.39$, nor the effect of EPI classification, $F_{(1.26)} = 1.23$, $p = 0.27$, were significant. The interaction between field dependence/independence and extraversion/introversion was also non-significant, $F_{(1.26)} = 1.56$, $p = 0.22$.

Despite insignificant interactions, Figure 2 shows that field-dependent and introvert participants answer more normal questions correctly that any other group. Extraverts in both groups and field-independent introverts perform on similar levels.
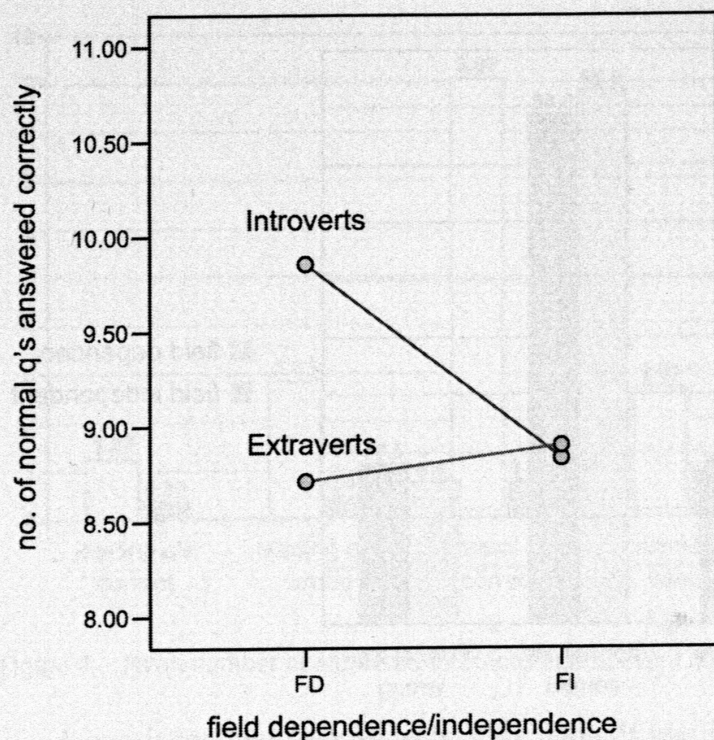
Figure 2. Performance of introverts and extraverts on number of correct answers to normal questions and FDI

Pearson's Product Moment Correlation Coefficient was used to investigate further relationships between field dependence personality traits and suggestibility to leading questions. The aim was also to assess the relationship between confidence and other variables.

There was a moderate positive relationship between the number of leading questions answered correctly and field dependence/independence ($r = 0.39$, $p = 0.03$) (see: *Appendix 2*).

Examining FD and FI subjects further highlights this finding (see Figure 3). These two groups do not differ significantly on any other type of questions, apart from the leading questions.
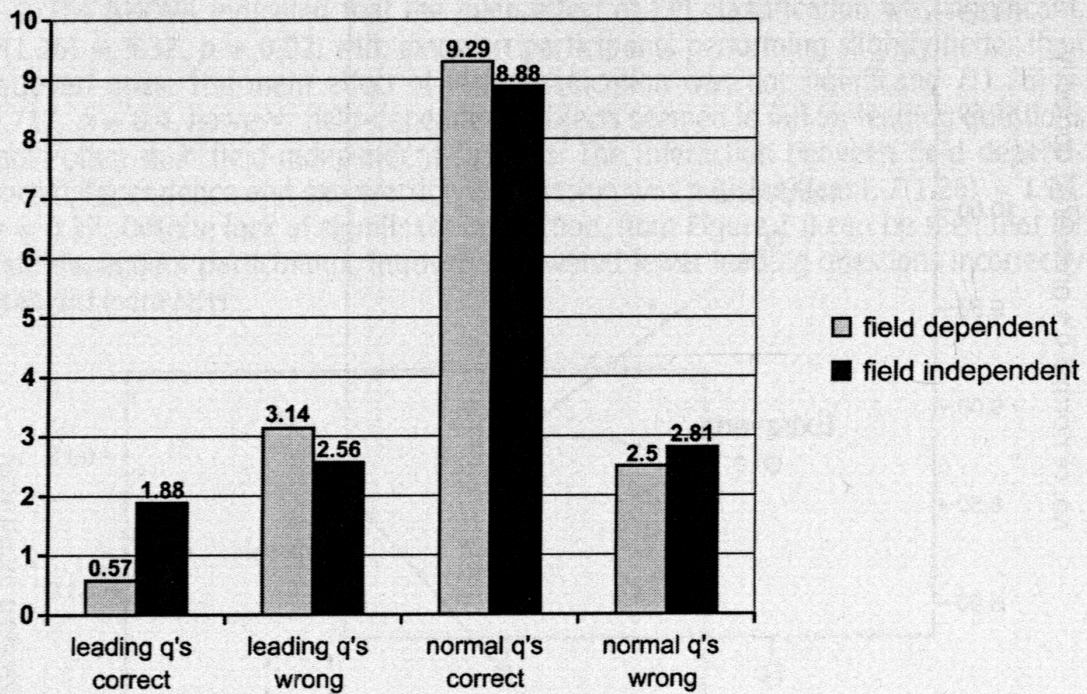
Figure 3. Mean number of questions answered correctly and incorrectly

Several correlations suggest that certain personality traits may be influencing suggestibility:
- extraversion scale and leading questions answered correctly correlated positively at $r = 0.41$, $p = 0.02$, also there was a positive correlation between extraversion and confidence for those questions ($r = 0.52$, $p = 0.003$)
- extraversion scale and leading questions answered incorrectly correlated negatively at $r = -0.5$, $p = 0.005$
- neuroticism scale correlated with number of wrong answers to leading questions at $r = 0.36$, $p = 0.04$

A connection between suggestibility and confidence was established by the following correlations:
- reported confidence and number of leading questions answered correctly correlated positively, $r = 0.58$, $p = 0.001$
- reported confidence for leading questions answered incorrectly was also positive, $r = 0.44$, $p = 0.01$

Figure 4 presents the mean number of questions answered in every category of questions and related mean confidence.
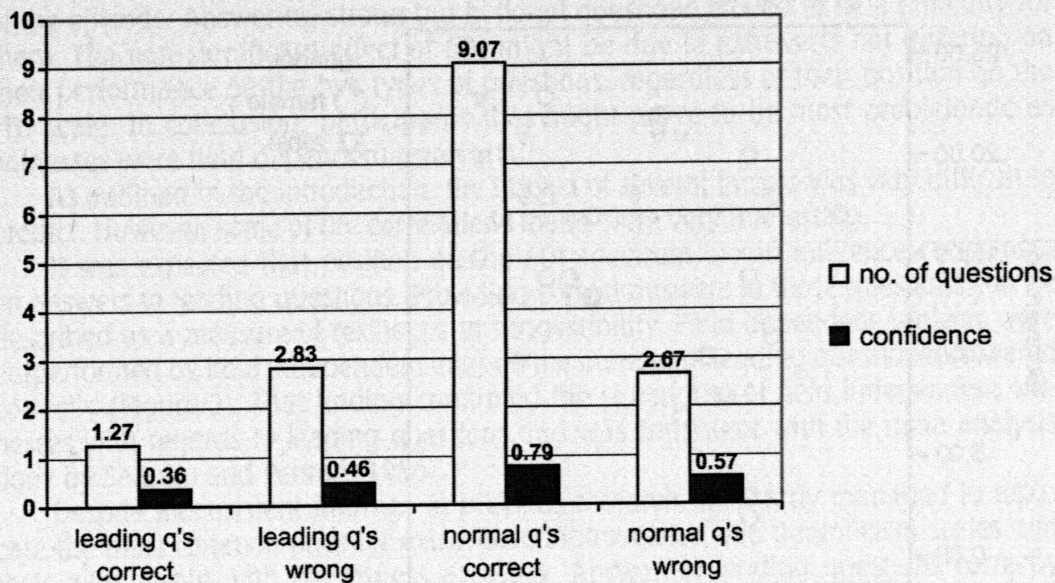
Figure 4. Mean number of answered questions and relevant confidence

A correlation that complements other findings was found between extraversion and FDI. Both factors have been found correlating with suggestibility and they also correlated with each other at a moderate level ($r = 0.38$, $p = 0.03$).

Additionally, the effect of gender was investigated by comparing performance of females and males on EFT, extraversion scale, correct number of leading questions answered and related confidence (Table 2).

Table 2. Performance of males and females on EFT, extraversion scale, correct number of leading questions answered and related confidence.

Gender	EFT	Extraversion/ introversion	Leading q's correct	Confidence
Female	10.00	13.84	1.16	0.31
Male	14.55	18.55	1.44	0.46

Further examination of gender distribution (Figure 5) revealed that whereas women's scores were spread evenly across extraversion scale and field dependence/ independence spectrum, majority of men were highly field-independent and more extravert.

strona 193
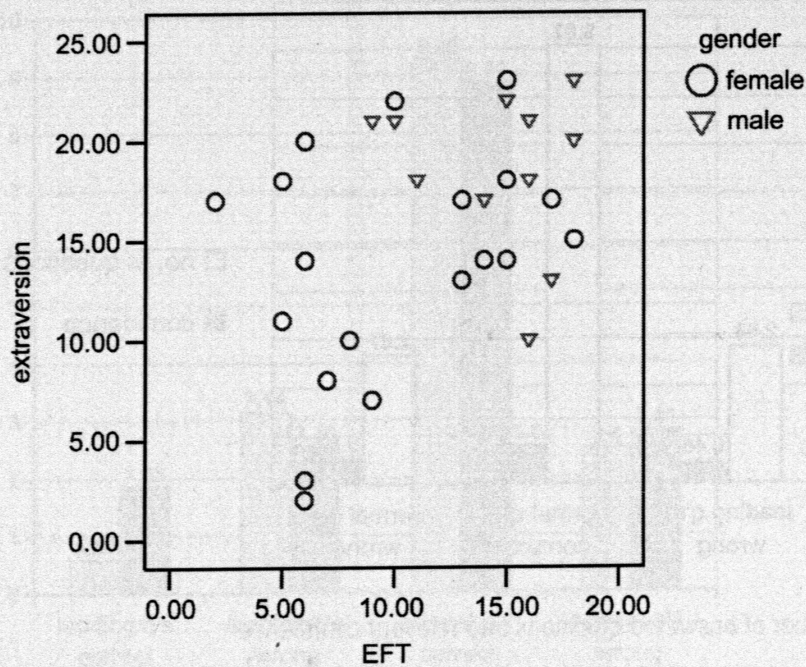
Figure 5. Distribution of males and females on EFT and extraversion/introversion scale

Discussion

The purpose of this study was to investigate the relationships between various characteristics of the witness and their overall impact on eyewitness testimony. Participants were assessed on their position on two dimensions: extraversion/introversion and field dependence/independence. It was predicted that these two factors would be connected to participants' ability to fall for leading questions. An additional variable – post-event confidence – was introduced to further analyse the effect of suggestibility, especially with different types of questions.

The main hypothesis was confirmed by the analysis of variance. The results of the present study show a difference in levels of susceptibility to leading questions. Introvert participants were significantly more predisposed to fall for leading questions as opposed to extravert subjects. This finding was in accordance with others, that found extraverts to be more resilient to misleading information (e.g. Gudjonsson et al., 2004). The main effect of the FDI dimension was found to be non-significant. However, further analysis confirmed previous findings that field independent participants tend to perform better and fall for leading questions less often than field dependent participants (e.g. Shapiro, Penrod, 1986). The importance and effect of leading questions seems to be even more accentuated by the results of the second ANOVA. Field dependent introverts actually outperformed extraverts answering more normal questions correctly. The advantage was not big but it shows that introvert participants had a good memory of the event and remembered many details

strona 194

of the episode. Answering strong but fictional questions proved to be a difficulty for them. The non-significant effect of FDI might be due to extraverts not differing on their performance on the two types of questions, regardless of their position on the FDI scale. In conclusion, participants that might prove to be most problematic as witnesses were field dependent introverts.

As outlined in the introduction, the impact of several factors was very difficult to predict. However, some of the correlations found were very interesting.

It was expected that position on the FDI spectrum would influence correctness on answers to leading questions. Providing correct answers to those questions can be described as a measure of resilience to suggestibility. Field dependent subjects were outperformed by field independent ones on the number of leading questions answered correctly (Figure 2). That finding confirmed the superiority of field independent witnesses with regards to leading questions and was consistent with the meta-analysis done by Shapiro and Penrod (1986).

Despite inconsistent findings in previous research, this study managed to replicate the most common ones for extraversion/introversion and neuroticism scales and their relationship with eyewitness accuracy. Answering leading questions correctly correlated positively with extraversion whereas incorrect answers correlated negatively. Although the confidence factor will be discussed in detail further on, it is important to notice that confidence for leading questions answered correctly was significantly higher in extraverts, that concurring with findings by Trouve and Libkuman (1992). The results also showed that participants high on the neuroticism scale got more leading questions wrong than those with low scores. Both results concur with findings of Gudjonsson (1983) and Gudjonsson et al. (2004). Thus, correlations also clearly indicate that extravert, emotionally stable witnesses would be more likely to provide more truthful evidence.

As pointed out by Bothwell, Deffenbacher, and Brigham (1987), and Kebbel and Gilchrist (2004), the confidence-accuracy relationship is not particularly strong. Meta-analysis of almost 40 separate tests of the post-hoc confidence and accuracy correlation found the average correlation to be 0.25 (Bothwell et al., 1987). Leippe (1980) suggested that this relationship is weak because there are factors that affect identification accuracy without influencing confidence (e.g. reconstructive processes in memory). The opposite can also happen, i.e., social influence processes would only affects confidence. This line of reasoning could maybe explain the surprising results of this study on confidence. It was found that participants who answered leading questions incorrectly were almost as confident as those who provided correct answers. The most possible explanation is that confidence was somehow influenced despite difference in actual knowledge. Perhaps the very form and power of leading questions elevated levels of confidence. Strong suggestibility of the questions made the subjects feel very strongly about their answers once they have provided one. Such a finding is very difficult to explain and would require further investigation. The question however remains about the practical importance of the whole problem, since witnesses confidence about their answers is generally a weak predictor of accuracy (Narby et al., 1996) and implications of it for the justice system seem to be limited.

The examination of gender differences revealed a slight superiority for male subjects. There have been consistent reports about differences in gender in relation to performance on EFT. Many researchers found that males tend to perform better than females (e.g. Huss, Kayson, 1985; Baron-Cohen, 1998). The effect of gender was yet again confirmed in present study. However, males also seemed to score higher on the extraversion scale. It has therefore been difficult to determine which dimension was responsible for the higher accuracy of males on leading questions and related confidence.

Several relationships between various factors and accuracy of a witness, reported in the literature, have been reproduced here. The present study established that leading questions influenced accuracy of the witness and that the relationship with confidence was distorted by the type of questions that were asked. Personality traits and cognitive style were shown to be reliable sources of prediction about witnesses abilities to deliver accurate testimony.

However, despite growing number of laboratory researches, the usefulness of these findings for the criminal justice system, especially with regards to suggestibility, can be arguable. In real cases particular evidence might be of a great importance for the prosecution as well as defence and the witness might not necessarily be questioned using confusing techniques. Furthermore, the use of those techniques can be very effectively minimised by appropriate intervention of the judge, especially in case of a very vulnerable witness (O'Kelly, Kebbell, Hatton, Johnson, 2003). For that reason it is of great importance to conduct experiments that replicate real life as closely as possible. It might have been an improvement of this study for example to randomise questions asked after the movie, which might have resulted in more explicit findings. It is also acknowledged that the tool used to test suggestibility have not been tested for validity before the research.

There is no doubt that information about a personality profile and a cognitive style of a witness may help to create an efficient environment and obtain the most accurate evidence. Psychologists are able to provide lawyers and other professionals with this valuable knowledge about the witness, and findings similar to those replicated in this study have made a significant contribution to the criminal justice system.

Acknowledgements

References

Baron-Cohen, S. (1998). Superiority on the Embedded Figures Task in autism and in normal males: evidence of an "innate talent"? *Behavioural and Brain Sciences, 21*, 408-409.

Bekerian, D. A., & Bowers, J. M. (1983). Eyewitness testimony: Were we mislead? *Journal of Experimental Psychology: Learning, Memory and Cognition, 9*, 139-145.

Berger, J. D., & Herringer, L. G. (1991). Individual Differences in Eyewitness Accuracy. *Journal of Social Psychology, 131*, 807-813.

Bothwell, R. K., Deffenbacher, K. A., & Brigham, J. C. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology, 72*, 691-695.

Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior, 14*, 41-55.

Emmet, D., Clifford, B. R., & Gwyer, P. (2003). An investigation of the interaction between cognitive style and context reinstatement on the memory performance of eyewitness. *Personality and Individual Differences, 34*, 1495-1508.

Endres, J. (1997). The Suggestibility of the Child Witness: The Role of Individual Differences and Their Assessment. *The Journal of Credibility Assessment and Witness Psychology, 1*, 44-67.

Eysenck, H. J., & Eysenck, S. B. (1991). *Manual of the Eysenck Personality Scales (EPS Adult).* London: Hodder & Stoughton Educational.

Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology, 9*, 215-227.

Gudjonsson, G. H., (1983). Suggestibility, intelligence, memory recall and personality: an experimental study. *The British Journal of Psychiatry, 142*, 35-37.

Gudjonsson, G. H., (1984b). A new scale of interrogative suggestibility. *Personality and Individual Differences, 5*, 303-314.

Gudjonsson, G. H., Sigurdson, J. F., Bragason, O. O., Einarsson, E., & Vladimarsdottir, E. B., (2004). Compliance and personality: the vulnerability of the unstable introvert. *European Journal of Personality, 18*, 435-443.

Hafstad, G. F., Memon, A., & Logie, R. (2004). Post-identification feedback and recollections of witnessing conditions in child witness. *Applied Cognitive Psychology, 18*, 901-912.

Huss, E. T., & Kayson, W. A., (1985) Effects of Age and Sex On Speed of Finding Embedded Figures. *Perceptual and Motor Skills, 61*, 591-594.

Ihelbaek, C., Love, T., Eilertsen, D. E., & Magnussen, S. (2003) Memory for a staged criminal event witnessed live and on video. *Memory, 11*, 319-327.

Kebbel, M. R., & Gilchrist, M. L. (2004) Eliciting evidence from eyewitnesses in court. In: J. R. Adler (Ed.) *Forensic Psychology: Concepts, debates and practise.* Devon: William Publishing.

Kebbel, M. R., & Johnson S. D. (2000). Lawyers' Questioning: The Effect of Confusing Questions on Witness Confidence and Accuracy. *Law and Human Behavior, 24*, 629-641.

Lane, S. M. Dividing Attention During a Witnessed Event Increases Eyewitness Suggestibility. *Applied Cognitive Psychology.* In press.

Leippe, M. R. (1980). Effects of integrative memorial and cognitive processes on the correspondence of eyewitness accuracy and confidence. *Law and Human Behavior, 4*, 261-273.

Loftus, E. F. (1975). Leading questions and eyewitness report. *Cognitive Psychology, 7*, 560-572.

Loftus, E. F. & Zanni, G. (1975). Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society, 5*, 86-88.

Narby, D. J., Cutler, B. L., & Penrod, S. (1996). The effects of witness, target, and situational factors on eyewitness identifications. In: S. L. Sporer, R. Malpass & G. Koehnken (Eds.) *Psychological Issues in Eyewitness Identification.* NJ: Mahwah, Lawrence Erlbaum.

O'Kelly, C. M. E., Kebbell, M. R., Hatton, C., & Johnson, S. D. (2003). When do Judges intervene in cases involving people with learning disabilities? *Legal and Criminological Psychology, 8*, 229-240.

Pezdek, K., & Greene, J. (1993). Testing Eyewitness Memory: Developing a Measure That Is More Resistant to Suggestibility. *Law and Human Behavior, 17*, 361-369.

Pozzulo, J. D., Coplan, R. J., & Wilson, J. (2005). Not to 'shy' to help catch a thief: Recall memory of shy versus non-shy witnesses. *Personality and individual differences, 39*, 1361-1372.

Schaefer, P. D., & Thomas, J. (1998). Difficulty of a Spatial Task and Sex Differences in Gains From Practice. *Perceptual and Motor Skills, 87*, 56-58.

Shaw, J. S., &, McClure, K. A. (1996). Repeated Postevent Questioning Can Lead to Elevated Levels of Eyewitness Confidence. *Law and Human Behavior, 20*, 629-653.

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

Skolnick, P., & Shaw, J. I. (2001). A comparison of eyewitness and physical evidence on mock-juror decision making. *Criminal Justice and Behavior, 28*, 614-630.

Smith, S. M. & Rothkopf, E. Z. (1984). Contextual Enrichment and Distribution of Practice in the Classroom. *Cognition and Instruction, 3*, 341-358.

Trouvé, R. J., & Libkuman, T. M. (1992). Eyewitness performance of personality types as a function of induced arousal. *America Journal of Psychology, 105*, 417-433.

Ward, R. A., 7 Loftus, E. F. (1985). Eyewitness performance in different psychological types. *The Journal of General Psychology, 112*, 1546-1557.

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

Wells, G. L., Lindsay, R. C. L., & Ferguson, T. J. (1979). Accuracy, confidence and juror perception in eyewitness identification. *Journal of Applied Psychology, 64*, 440-448.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22*, 603-645.

Witkin, H., Dyk, R. B., Faterson, H. F., & Karp, S. A., (1971). *A manual for the Embedded Figures Test.*, Palo Alto: Consulting Psychologist Press.

Appendix 1

**Please answer all the questions,
according to what you remember from the video**

5 – very confident
4 – quite confident
3 – unsure
2 – slightly unconfident
1 – not confident at all

1. What was the employee stocking the shelves with at the beginning of the video?
2. What was the customer wearing?
3. What colour was the doorframe?
4. Which of the robbers entered the shop first?
5. Did the robbers steal a video from the 'action' or 'new releases' section?
6. How did the robbers disguise their faces?
7. How would you describe the cashier's hairstyle?
8. Did the cashier open the till on his left or right first?
9. What type of bag was handed to the cashier?
10. Did the robber hand the bag to the cashier with his left or right hand?
11. How did the main robber get the employee to hurry up?
12. How would you describe the robber's accent?
13. With which hand did the robber by the door point the gun at the cashier?
14. What colour hair did the robber by the door have?
15. Was the price offer on the poster on the door £4.99 or £5.99?
16. Which of the robbers pushed the customer?
17. Did the main robber hand over the bag with money to his accomplice before or after taking the face cover?
18. In which direction did the robbers run off after leaving the shop?

Table of correlations

		AGE	GEFT	LQ_CORRE	CONFIDEN	LOWRONG	CONFID_A	LQ_NO_AN	NQCORREC	CONFID_B	NQWRONG	CONFID_C	NQ_NO_AN	PSYCHOTI	EXTRAVER	NEUROTIC
AGE	Pearson Correlation	1.000	-.185	.000	-.114	.022	-.104	-.014	-.068	-.212	.094	.046	-.020	-.151	-.260	-.039
	Sig. (2-tailed)		.327	.998	.549	.908	.584	.943	.723	.261	.621	.809	.916	.424	.165	.836
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
GEFT	Pearson Correlation	-.185	1.000	.396*	.403*	-.222	.085	-.238	-.079	-.017	.009	.102	.044	.364*	.380*	-.064
	Sig. (2-tailed)	.327		.030	.027	.238	.655	.206	.678	.927	.961	.591	.819	.048	.038	.735
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
LQ_CORRE	Pearson Correlation	.000	.396*	1.000	.584**	-.759**	-.283	-.256	-.319	.227	.290	.028	.002	.108	.412*	-.179
	Sig. (2-tailed)	.998	.030		.001	.000	.130	.171	.086	.228	.120	.882	.992	.569	.024	.343
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
CONFIDEN	Pearson Correlation	-.114	.403*	.584**	1.000	-.569**	.001	.039	-.149	.184	.038	.110	.080	.337	.520**	-.267
	Sig. (2-tailed)	.549	.027	.001		.001	.995	.840	.433	.331	.844	.562	.673	.069	.003	.153
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
LOWRONG	Pearson Correlation	.022	-.222	-.759**	-.569**	1.000	.440*	-.427*	.505**	-.166	-.039	.147	-.488**	-.246	-.504**	.365*
	Sig. (2-tailed)	.908	.238	.000	.001		.015	.019	.004	.382	.839	.437	.006	.190	.005	.047
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
CONFID_A	Pearson Correlation	-.104	.085	-.283	.001	.440*	1.000	-.336	.380*	-.060	.084	.279	-.317	-.235	-.037	.165
	Sig. (2-tailed)	.584	.655	.130	.995	.015		.069	.038	.751	.658	.135	.088	.212	.847	.383
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
LQ_NO_AN	Pearson Correlation	-.014	-.238	-.256	.039	-.427*	-.336	1.000	-.295	-.039	-.500**	-.219	.685**	.258	.184	-.318
	Sig. (2-tailed)	.943	.206	.171	.840	.019	.069		.114	.839	.005	.245	.000	.169	.331	.087
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
NQCORREC	Pearson Correlation	-.068	-.079	-.319	-.149	.505**	.380*	-.295	1.000	.122	-.311	.152	-.586**	-.250	-.233	.203
	Sig. (2-tailed)	.723	.678	.086	.433	.004	.038	.114		.521	.094	.422	.001	.183	.216	.281
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
CONFID_B	Pearson Correlation	-.212	-.017	.227	.184	-.166	-.060	-.039	.122	1.000	.084	.201	-.193	.146	.198	-.080
	Sig. (2-tailed)	.261	.927	.228	.331	.382	.751	.839	.521		.659	.287	.307	.442	.294	.674
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
NQWRONG	Pearson Correlation	.094	.009	.290	.038	-.039	.084	-.500**	-.311	.084	1.000	.311	-.579**	-.093	-.105	.092
	Sig. (2-tailed)	.621	.961	.120	.844	.658	.659	.005	.094	.659		.095	.001	.624	.581	.627
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
CONFID_C	Pearson Correlation	.046	.102	.028	.110	.147	.279	-.219	.152	.201	.311	1.000	-.378*	-.133	.045	.119
	Sig. (2-tailed)	.809	.591	.882	.562	.437	.135	.245	.422	.287	.095		.039	.485	.813	.532
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
NQ_NO_AN	Pearson Correlation	-.020	.044	.002	.080	-.488**	-.317	.685**	-.586**	-.193	-.579**	-.378*	1.000	.262	.260	-.234
	Sig. (2-tailed)	.916	.819	.992	.673	.006	.088	.000	.001	.307	.001	.039		.162	.165	.213
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
PSYCHOTI	Pearson Correlation	-.151	.364*	.108	.337	-.246	-.235	.258	-.250	.146	-.093	-.133	.262	1.000	.499**	-.091
	Sig. (2-tailed)	.424	.048	.569	.069	.190	.212	.169	.183	.442	.624	.485	.162		.005	.634
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
EXTRAVER	Pearson Correlation	-.260	.380*	.412*	.520**	-.504**	-.037	.184	-.233	.198	-.105	.045	.260	.499**	1.000	-.130
	Sig. (2-tailed)	.165	.038	.024	.003	.005	.847	.331	.216	.294	.581	.813	.165	.005		.493
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
NEUROTIC	Pearson Correlation	-.039	-.064	-.179	-.267	.365*	.165	-.318	.203	-.080	.092	.119	-.234	-.091	-.130	1.000
	Sig. (2-tailed)	.836	.735	.343	.153	.047	.383	.087	.281	.674	.627	.532	.213	.634	.493	
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).