AN INVESTIGATION INTO PHYSIOGNOMIC PERCEPTION: ACCURATE JUDGMENTS OF PSYCHOPATHIC TRAITS AND CRIMINALITY FROM FACIAL PHOTOGRAPHS

A thesis
submitted in partial fulfilment
of the requirements for the Degree
of
Master of Science in Psychology
in the
University of Canterbury
by
R. F. Grundy

University of Canterbury
2000

My thanks to Drs. Maureen Barnes and Lucy Johnston, who supervised every stage in the production of this thesis. Your collective wisdom was invaluable and your patience never faltered. Thank you also to Drs. Grace and Owen, who gave their opinions regarding methods of data analysis.

Contents

List of Tables

List of Figures

Abstract

Research has shown that some aspects of personality can be accurately read from the permanent features of a person's face. This accuracy has been explained through reference to actual relationships between underlying traits and facial appearance. These could be caused through a self-fulfilling prophecy that arises from the effect of facial stereotypes on personality development, from habitual expressions leaving permanent traces on the face, or from common biological or environmental causes of both facial appearance and character. The latter two explanations are associated with physiognomy, the direct perception of underlying character through reading faces. An attempt was made to find support for the validity of physiognomy by controlling those physical characteristics of faces that have been linked to processes of over-generalisation, and thus might lead to self-fulfilling prophecies. Other potential cues to underlying character that were irrelevant to physiognomy were also controlled. It was predicted that participants would accurately perceive psychopathic traits and criminal propensity from facial photographs of sexual serial killers and non-criminals. This prediction was supported. However, the accuracy found may have resulted from a failure to control differences between the sexual serial killers and non-criminals relating to certain physical characteristics of the photographs and their subjects. These variables were found to correlate with participants' judgments regarding psychopathic traits and criminality. It was concluded that the results were consistent with processes of over-generalisation, but that the occurrence of direct physiognomic perception could not be ruled out. Suggestions were made regarding future research that might isolate the direct perception of character from accuracy due to self-fulfilling prophecies. The importance of controlling confounding features such as grooming and transient facial expression was emphasised.

Introduction

Physiognomy, or reading character from permanent features of the face, has a long history. References to this erstwhile field of study have been found in ancient Assyrian texts, while major proponents have arisen throughout the development of Western science; from Aristotle, to Lavatar in the eighteenth century, Lombroso in the late nineteenth century, and even up to Merton, whose physiognomic system held considerable sway among American businessmen in the first half of the twentieth century (Atwan, 1984; Brandt, 1980; Mainwaring, 1980). While physiognomy has shown such temporal persistence, so too have measures taken against it, including the Catholic church declaring it a heresy (Berry & Brownlow, 1989; Brandt, 1980), secular powers punishing its practise as a crime (Allport, 1937) and research psychologists (e.g. Brandenburg, 1926; Cleeton & Knight, 1924) resoundingly concluding that there is "no evidence that the structure of the face concomitantly varies with character, and much evidence that it does not" (Knight, 1932, p. 420).

More recently, investigations into the continued public belief in the validity of face reading (Berry & McArthur, 1986) have taken place under the rubric of stereotyping (McArthur, 1982). The emphasis has shifted away from determining the accuracy of physiognomy to the role of cognitive processes in developing and maintaining facial stereotypes (e.g. Laser & Mathie, 1982; McArthur, 1982; Nakdimen, 1984; Secord, 1958). It has become clear that there is a relative consensus to be found among the judgments of participants who are asked to infer a number of varied underlying characteristics from viewing static faces (e.g. Bull & Green, 1980; Goldstein, Chance, & Gilbert, 1984; Hochberg & Galper, 1974; Lasswell & Parshall, 1962; Mason, 1959; Nash, 1958; Roll & Verinis, 1971; Samuels, 1939; Sappenfield, 1977; Secord, Bevan, & Dukes, 1953; Secord, Dukes, & Bevan, 1954; Secord & Muthard, 1955a,b; Shoemaker, South, & Lowe, 1973; Yarmey, 1993). However, there are reasons, both theoretical and empirical, to believe that this consistency may not be entirely due to a common background of learned stereotypes. It might instead reflect an underlying kernel of truth in judgments from the fixed features of faces, a degree of accuracy in

person perception that would be consistent with the concept of physiognomy (Berry, 1990; Berry & Wero, 1993; McArthur, 1982).

Theoretical Expectations of Physiognomic Accuracy

In general, specific physiognomic systems provide little reason for an expectation of perceptual accuracy, either because they are implausible themselves or essentially atheoretical. For instance, the Aristotelian theory that a person's character reflects the attributes of the animal that he or she most resembles holds no persuasive force today, while Lavatar's extensive and influential physiognomic system was simply based on talents for measurement and drawing plus unsystematic extrapolation from the co-occurrence of his friends' facial features and their personalities (Brandt, 1980). Many later approaches not only lend no theoretical weight to accurate facial person perception; they actually presuppose the inaccuracy of facial judgments, seeking only to explain the empirical consensus between judges.

Secord (1958), for example, proposes a role for commonly held cultural factors and inference processes in organising facial impressions of personality. Culture is said to enhance the congruency of judgments by placing emphasis on certain facial cues (e.g. lipstick), through stereotypes (e.g. of race, age and sex) and through common meanings for emotional expressions. He argues that people use analogy to "bridge the gap from stimulus information to personality judgments" when asked to come to a conclusion based upon impoverished information (Secord, Stritch, & Johnson, 1960, p. 329). Secord (1958) lists five kinds of inference processes relevant to judgments from facial photographs. Temporal extension involves perceiving a transient characteristic as an enduring attribute, for example a temporarily frowning person is thought to be consistently grumpy. Parataxis entails generalising from a significant other, for example a physical resemblance to another being equated with a psychological similarity. In categorisation, facial cues are used to put a person into a category with which the observer associates certain personality attributes. For instance, an observer might perceive someone as old and hence responsible, due to a personal belief that older individuals are more responsible. Inference based on functional qualities involves such reasoning as that of thick lips suggesting high sexuality or a high forehead indicating intelligence. Meanwhile, metaphorical

generalisation is inference of a more abstract type, like coarse hair implying a coarse nature. Importantly, none of these processes is theoretically tied to a position implying any accuracy in facial perception. However, while more recent researchers have perpetuated this emphasis on cognitive processes that produce judgment consensus, some have grounded these processes in the context of functional, accurate perceptions.

Nakdimen's (1984) interesting discussion of the physiognomic basis of sexual stereotyping shares much in common with Secord's (1958) ideas relating to judgment consensus but the cues which are subject to erroneous generalisation are more explicitly related to aspects of human biology, which themselves might be expected to be accurately perceived. He introduces the concepts of "nonverbal quasi-communication" (p. 500), denoting anatomical configurations that mimic body language, for example emotional expression, and "nonverbal quasi-information" (p. 502), meant to cover appearances that provide false information unrelated to body language, for example regarding personality. External support for the quasi-communication of emotional expression can be found in a study where individual structural features in a face (e.g. the thickness of eyebrows) were manipulated, leading to changes in perceived expression (Laser & Mathie, 1982). The effects of quasi-information have been demonstrated through ratings of dominance being increased by the addition of mature facial features to Identi-Kit faces (Keating, 1985).

Nakdimen (1984) reasons that certain features in women's physical appearance have helped create sexual stereotypes relating to personality traits. For example, as women tend to have more arched eyebrows, and as raised eyebrows are components of expressions of surprise and of non-analytic attention, then this could contribute to the stereotype that women are more intuitive and less logical than men. The structurally high, arched eyebrows would quasi-communicate a non-analytic cognitive character. Furthermore, as lowered eyebrows are a sign of dominance in non-human primates, arched brows might contribute to stereotypes of submissiveness. Similarly, high cheekbones are argued to quasi-communicate smiling and a delicate, moderately concave bridge to the nose, as a sign of an infant's immaturity, could provide quasi-information relating to helplessness and other immature personality traits. Whether or not there is great merit in his thesis that the sexual dimorphism of anatomy fosters the

development of stereotypical beliefs regarding personality during childhood, Nakdimen's (1984) comments emphasise a link between facial stereotypes and evolutionarily adaptive perceptions of expression and maturity.

Zebrowitz (1997) details a more comprehensive model to explain accurate consensus in facial judgments by reference to real appearance-trait relationships (see also Zebrowitz & Collins, 1997). She acknowledges the possibility that both biology and environment could constitute a common cause for facial appearance and psychological traits that would be open to physiognomic perception but places greatest emphasis on putative causal links between facial appearance and psychological traits. Zebrowitz (1997) suggests that personality could affect resting facial appearance through either a "Dorian Gray" effect, whereby personality moulds the face into its own likeness, or an artifice effect, where an appearance incongruent with personality is produced. A Dorian Gray effect could involve, for instance, pervasive anger affecting jaw development through muscle tension or a person with criminal tendencies choosing to undergo facial tattooing. An example of an artifice effect would be a dishonest woman who rejects the use of make-up in order to appear naïve. In turn, facial appearance could affect personality indirectly through manipulation of the environment. This might produce a self-fulfilling prophecy effect, for example, an ugly child who is treated badly due to his or her appearance responds by growing up with an "ugly" personality. In the alternative, a self-defeating prophecy effect might occur, where a person's character becomes less stereotypical through compensating for the environment brought about by his or her facial features. For example, a babyfaced male could become more aggressive to prove his maturity.

Like Secord (1958) and Nakdimen (1984), Zebrowitz (1997) reserves an important place for processes of over-generalisation in facial perception. She sees over-generalisation as the foundation of the continuing practice of reading traits from faces. Zebrowitz (1997) reviews evidence that age, sex, ethnicity/race, identity, emotion and fitness can all, to varying extents, be judged accurately from the face, and that some of these characteristics are perceived through specialised neural mechanisms. These are presumably evolutionarily adaptive. She takes the position that the evolutionary importance of detecting these qualities might have led to "such a strong preparedness

to respond to the facial qualities that reveal these attributes that our responses are over-generalized to individuals whose appearance merely resembles them" (p. 56). This is essentially the same stance as that held by Nakdimen (1984). Thus some of the congruency between facial judgments is explained through a functional (ecological) approach to person perception, which is placed within the context of evolutionary theory.

The ecological approach to perception is based upon Gibson's (1979) theory of affordances. An affordance is an opportunity to act, interact, or be acted upon that is provided by the environment. What the environment offers to an animal is hypothesised to be directly perceived through extracting invariant combinations of variables from the stimulus flux. Thus the ecological approach to person perception emphasises that socially important qualities actually structure stimulus information and can be directly perceived, without recourse to probabilistic inference (Zebrowitz & Collins, 1997). For instance, the fixed features of an individual's face comprise invariants that could lead to the direct perception of antisocial tendencies, where "antisocial" can be seen as a descriptive label for what is afforded to another through social interaction with that individual. Such direct perception would not rely upon cognitive processes weighing the probability of being antisocial against actuarial knowledge, or stereotypes, that link criminal behaviour to qualities such as age, race or grooming. Hence, an ecological approach suggests that accurate consensus in facial judgments need not necessarily follow from the influence of over-generalisation, although this could have an impact on actual appearance-trait relationships, but might instead involve true physiognomy.

It should be pointed out that, while it provides a framework for direct perception from structural invariants, an ecological perspective is not entirely congruent with physiognomy. This perspective also suggests that information is typically revealed in dynamic events rather than static displays (McArthur & Baron, 1983) because "perception of the environment is not based on a sequence of snapshots but on invariant-extraction from a flux" (Gibson, 1979, p. 304). It follows that the informative value of the resting face would be less than that gleaned from observation having access to additional aspects of appearance and behaviour. Nonetheless, the

ecological perspective on person perception provides a theoretical reason to expect valid physiognomy in relation to certain characteristics.

More specifically, this approach predicts that qualities in others will be accurately perceived if such perception would encourage behaviour that is adaptive for the self or for the species in general (Berry & Wero, 1993; McArthur & Baron, 1983). For example, antisocial tendencies or physical aggressiveness should be perceivable, as such perception would enable efforts to protect oneself from harm that would promote survival at both an individual and species level. Thus the ecological approach assumes that there is a hierarchy in how readily social properties will be perceived (McArthur & Baron, 1983). As perception is intrinsically linked to adaptive behaviour, perceptions of affordances should be more accurate than perceptions of qualities that do not afford opportunities to act adaptively.

In other words, we should be able to directly see whether someone is dangerous or sexually receptive from structural (e.g. fixed facial structure) or transformational (e.g. the movements of emotional expressions) invariants because this will have an impact on what we can do to enhance our survivability. There would be less reason to be able to perceive someone's abilities of mathematical reasoning or aesthetic appreciation. If any personality traits can be accurately perceived through physiognomy then, according to the ecological approach, they should include those that impact the most strongly upon those social interactions necessary for survival, such as finding a mate, parenting and working co-operatively. Investigations into such traits might allow the isolation of that part of accurate consensus in facial judgments which is not due to the effects of over-generalisation.

Empirical Evidence Supporting Physiognomy

The feasibility of accurate physiognomic perception as an explanation for judgment consensus is supported by data that indicate that not every aspect of the consensus identified as arising from facial stereotypes is necessarily learned. Berry and Wero (1993) argue that findings of inter-cultural homogeneity in linking inferred personality to physiognomic features (e.g. McArthur & Berry, 1987; Secord & Bevan, 1956) run contrary to facial judgments being based entirely upon a socially

constructed belief system. Further evidence pointing to a non-learned aspect of facial judgments includes the observation that children as young as 2-3 months will look at an attractive facial stimulus (as determined by the consensus of adult judges) in preference to an unattractive one. Presumably, they can therefore discriminate some factor inherent to facial attractiveness before being exposed to cultural stereotypes of beauty (Langlois et al., 1987). Certainly, Dion (1973) found that by only three years of age children communicate consistent judgments of peer facial attractiveness and identify physically attractive peers as both preferred for friendship and less likely to engage in antisocial activity.

In addition to the above evidence hinting at a non-learned basis to what has been identified as facial stereotyping, it could be suggested that the proof discrediting physiognomy is not as strong as is typically assumed. Instead, it is restricted to certain types of physiognomic perception. The various approaches to physiognomy can be characterised into two main categories. The first entails the resemblance of a face to someone or something else combined with an assumption that similarity in appearance implies a similarity in psychological characteristics. For example, all people with similar appearances due to their small chins are seen as weak, or those with flat, leonine faces are perceived as brave. The second involves looking for "traces of... expressive facial postures as clues to the common emotional and cognitive states of an individual" (Alley, 1988, p. 169, original emphasis removed). These different versions of physiognomy have been named "structural" and "expressional" respectively (Knight, 1932). This distinction has parallels in rough divisions between, on the one side, bone structure, claims regarding the meaning of particular facial features, permanence of features and inherited characteristics, and, on the other, musculature, judgments based upon an overall impression from viewing the whole face, malleability of features and environmental effects.

Unequivocal data questioning the validity of physiognomy appear to be restricted to those studies investigating the relationship between character traits and the measurement of particular facial features, that is structural, feature-based physiognomy. For example, Cleeton and Knight (1924) investigated relationships between groups of physical measurements taken from the face (e.g. length of nose, distance between eyes, width of the head) and eight character traits that were assessed

by close associates (sound judgment, intellectual capacity, frankness, will power, ability to make friends, leadership, originality, and impulsiveness). The resulting distribution of 201 correlation coefficients gave no indication of more than a chance connection between specific facial measurements and personality. Likewise, after reviewing studies of physiognomy from early in the 20th century, Brandenburg (1926) concluded that "not one... has produced any evidence favorable to the physiognomic system" (p. 588). Again, this was not based on studies relating to expressional physiognomy or which involved global facial impressions but upon investigations into traditional physiognomic claims regarding physical characteristics such as the colour of complexion and hair, facial profile, the texture and consistency of the hair and skin, peculiarities of the eyes, and characteristics of the mouth.

Furthermore, the degree of certainty engendered by those early studies discrediting traditional, structural physiognomy may have been inappropriate, not only because it served to taint *any* suggestion that personality can be read from the face but also because some later evidence suggests that specific facial features actually do correlate with personality. Squier and Mew (1981) compared 40 participants who were categorised as having either long or short faces through exacting measurements taken from radiographic pictures of the skeletal structure of the face and jaw. Each participant completed a self-report measure of personality. The researchers found that "subjects with long, angular faces were... more responsive, assertive, and genuine than subjects with short, square faces who were more restrained, conforming, and shrewd" (p. 151). This conclusion is actually supportive of the general thrust, if not the details, of the structural, feature-based physiognomic systems. It also corroborates McCabe's (1928) reanalysis of an earlier study, where the measured convexity of the lower faces of 24 female members of a university sorority was related to their rankings of each other according to the character traits of optimism, activity, ambition, will power, domination and popularity. Thus it seems possible that the structure of the jaw and lower face is truly related to personality. However, these results are as equally consistent with an appearance-trait correlation occurring due to the effects of self-fulfilling prophecies as they are with those common biological causes that are associated with structural physiognomy.

The evidence regarding the validity of global facial judgments, those made without reference to specific facial features or the tenets of popularised systems of physiognomy, is probably best described as mixed, and is certainly not deserving of the castigation given to traditional physiognomy. On one hand, studies investigating the perception of leadership (Mason, 1959) and intelligence (e.g. Cook, 1939; Gurnee, 1934; Ray, 1958) have failed to provide any real support for whole-face physiognomy. On the other hand, McCabe (1928) reports results from two studies that appear to advance a perceivable relationship between character and global facial appearance. Both studies used the same method, whereby members of a sorority with personal knowledge of their fellows ranked them on the traits of neatness, conceit, sociability, humour, likability (sic), intelligence, refinement, beauty, snobbishness and vulgarity. An equal number of strangers then ranked the same people on those traits according to photographs revealing the head, neck and shoulders. Disregarding beauty, which is an overtly physical dimension, the measured traits revealed a consistent pattern of positive correlation between the pooled judgments of acquainted and stranger judges. While no particular traits were consistently perceived with great accuracy across the two studies, the overall trend of positive correlation indicated some ability to judge character from the face.

Additional research has also supported the validity of global facial judgments in relation to specific traits or characteristics. Statistical links have been forged between perceived (from global assessments of facial photographs) and actual (from self-ratings or objective criteria) measures of: the power and warmth of personality (Berry, 1990, 1991), honesty (Berry, 1990; Bond, Berry, & Omar, 1994; Zebrowitz, Voinescu, & Collins, 1996), the particular offences perpetrated by criminals (Kozeny, 1962 cited by Bull & Green, 1980; Thornton, 1939), Machiavellianism (Cherulnik, Way, Ames, & Hutto, 1981) and dominance/leadership (Cherulnik, Turns, & Wilderman, 1990; Mazur, Mazur, & Keating, 1984). Taken as a whole, these results provide the strongest argument for accurate facial person perception, although they do not differentiate between explanations for that accuracy, whether valid physiognomy or the self-fulfilling nature of over-generalisations. These results also hint at the intriguing possibility that those traits that have been accurately judged from the face might be distinguished upon theoretical grounds from those that cannot be perceived.

An initial scrutiny might suggest that the characteristics listed above share a social, interpersonal quality. In contrast, intelligence, which has often proven ineluctable to physiognomic judgments (e.g. Cook, 1939; Gurnee, 1934; Ray, 1958), might be better thought of as instrumental in the performance of acts rather than strongly predictive of the social desirability of choices of action. Such a distinction would fit well with the prediction from ecological theory that affordances should be more accurately perceived than those personality traits which do not afford adaptive action (Zebrowitz & Collins, 1997), and hence also the possibility of direct physiognomic perception. It remains to be seen if further data will support this nascent pattern.

Finally, complementing the empirical links between whole-face judgments and particular aspects of character is the evidence relating global *descriptions* of appearance to personality. Such links between facial appearance and personality could be the bases of accurate physiognomic perception. That is, facial features that contribute to a physically attractive appearance might also directly convey personality to an observer making a physiognomic judgment. These global descriptions of appearance have also been related to stereotypes and thus form a natural area for investigating the influence of appearance stereotypes on the development of personality. Two widely researched facial stereotypes are those of the physically attractive, who are seen as sociable, socially skilled, popular, dominant and confident (Feingold, 1992), and those who are babyfaced, perceived as warm, submissive, naïve and less antisocial in their behaviour (Zebrowitz, Collins, & Dutta, 1998). These broad facial descriptions have also been found to correlate with actual personality and functioning, although not overwhelmingly so (see Berry & Wero, 1993; Berscheid & Walster, 1974; Bull & Rumsey, 1988; Sorell & Nowak, 1981; Zebrowitz, 1997). For example, male babyfaceness has been related to self-rated approachability and warmth (Berry & Brownlow, 1989) while its presence in childhood has actually predicted a less prototypically babyfaced personality in adolescents (Zebrowitz, Andreoletti, Collins, Lee, & Blumenthal, 1998; Zebrowitz, Collins, & Dutta, 1998). A meta-analytic review has linked attractiveness to actual popularity, sexual experience, social skill, decreased loneliness and lower social anxiety (Feingold, 1992), while individual studies have suggested advantages from academic success (Lerner & Lerner, 1977; Dickey-Bryant, Lautenschlager, Mendoza, & Abrahams, 1986) to enhanced psychiatric recovery (Farina, Austad, Burns, Bugglin, & Fischer,

1986). High earlier attractiveness has also recently been associated with a longitudinal lagged effect on sociability and hostility in participants during their fifties (Zebrowitz, Collins, & Dutta, 1998).

So, to summarise, there is evidence that observers can accurately estimate personality traits from assessments of fixed facial features, that some aspects of actual facial structure and broad facial descriptions like physical attractiveness are related to personality, and that attributions of personality from facial appearance are not necessarily learned. These pieces of evidence support valid physiognomy. However, as emphasised by Zebrowitz's (1997) model of accurate consensus, accurate facial judgments might also be explained in reference to the effects of over-generalisation and self-fulfilling prophecies due to facial stereotypes. The previous research, while demonstrating instances in which participants accurately rated traits from viewing static faces, does not really identify exactly why their judgments were accurate. The literature cannot be said to strongly support any one of the causal explanations of over-generalisation, expressional physiognomy or structural physiognomy over the others. This lack of specificity regarding the cause of accuracy appears to be due to a difficulty in isolating aspects of the face that cannot play a part in both over-generalisation and physiognomy.

Competing Explanations for Accuracy in Facial Judgments

Researchers provide four main explanations for the apparent kernel of truth found in certain judgments made from the fixed features of faces. These explanations can be placed within Zebrowitz's (1997) model of accurate consensus but are organised more specifically around the nature of the causes of accuracy, whether due to stereotypes, traces of expression, an underlying biological cause, or other factors that are not traditionally associated with physiognomy. The first is the concept of the self-fulfilling prophecy, or that facial stereotypes are the causes of their own confirmation by partially determining the environment in which people live and hence their personality development in reaction to that environment (Zebrowitz, Collins, & Dutta, 1998). There is considerable experimental evidence in support of the proposed means by which self-fulfilling prophecies develop; most importantly behavioural confirmation, differential expectations and differential treatment (Adams, 1977;

Langlois, 1986; Langlois & Stephan, 1981; Lerner & Lerner, 1977; Snyder, 1984; Sorrell & Nowak, 1981).

In the case of self-fulfilling prophecies the correlation between actual personality and attributions of personality based on facial appearance cannot be described as physiognomic, there being no indication of direct perception of an individual's underlying characteristics. For instance, a participant might attribute social confidence to all people with physically attractive faces and, because of an effect of attractiveness on their social environments which encourages sociability, this attribution may prove accurate to some extent. However, such accuracy would not imply an ability to read character directly from the face of an individual, only an ability to recognise physical attractiveness, which is indirectly related to a lack of social anxiety through the effect of stereotypical beliefs held about attractive people on their social environments. In other words, facial stereotypes might predict characteristics over a population, but no confidence could be held in facial assessments of any individual based upon these over-generalisations. Longitudinal research, providing evidence for the effects of physical attractiveness on later personality (Zebrowitz, Collins, & Dutta, 1998) supports a role for self-fulfilling prophecies in the accuracy of facial judgments, but cannot rule out the possibility of a shared biological cause for both attractiveness and personality. However, the research where links between early appearance and later personality have been shown to be dependent upon the stability of earlier appearance, is particularly consistent with a self-fulfilling prophecy causal effect (Zebrowitz & Collins, 1997). This is because it is generally thought that the cumulative effect of appearance on an individual's personality, which comprises a self-fulfilling prophecy, would require an appearance, and resulting environmental influences, that remain consistent over time (Adams, 1977; Berscheid & Walster, 1974; Langlois, 1986; Zebrowitz, Olson, & Hoffman, 1993).

Another possible route by which accurate predictions could be associated with facial assessments is through the direct expression of underlying character (Berry & Wero, 1993), which can be equated with expressional physiognomy but is a narrower concept than Zebrowitz's (1997) Dorian Gray effect. Here personality changes faces rather than faces affecting personality, and any accuracy in participants' judgments

should rely upon correctly reading facial traces of individuals' personalities. It is supposed that personality may mould facial bone structure through muscle tonus (Squier & Mew, 1981) and that musculature, flesh and facial expression reflect personality (Allport, 1937; Knight, 1932). Furthermore, findings of emotional asymmetry in the resting face support the contention that facial appearance can come to reflect its habitual expressions (personality). That is, the left hemiface, which is used more in both posed and spontaneous emotional expression (Borod, 1993), is perceived as more emotionally expressive even when it is at rest (McGee & Skinner, 1987), implying that the higher levels of emotional expression leave traces in the resting face. Similarly, the negative personality descriptions of cold and cruel, which might be linked to a lack of perceived emotion, have been ascribed more to right than left resting-hemiface composite portraits (Rhodes & Lynskey, 1990).

More convincingly, Malatesta, Fiore and Messina (1987) found that participants' errors in attributing emotion to the posed facial expressions of older people correlated with the actual emotional traits of those people. For instance, characteristically angry older people were more often misjudged to be angry when attempting to keep a neutral face or posing other expressions. Taken along with the literature relating to facial asymmetry, their data support the notion that habits of emotion and, by implication, personality leave permanent imprints on the face that might be validly perceived by others. However, it should be noted that the correlations between personality and fixed facial appearance that have been taken to support expressional physiognomy could have also resulted from a self-fulfilling prophecy effect or some common biological factor. Evidence of a Dorian Gray effect, whereby earlier personality is manifested in later facial appearance, has not been found in the two major longitudinal studies of appearance and personality (Zebrowitz, Collins, & Dutta, 1998; Zebrowitz et al., 1996); barring an influence that was attributable to the use of make-up.

The third major explanation for physiognomic accuracy is that a biological or genetic cause (Berry & Wero, 1993) common to both personality and facial features leads to the correlation between facial assessments and the predicted characteristic. This explanation is most closely associated with structural physiognomy. The concept is clearly exemplified in genetic defects such as Down's syndrome where intelligence

and face are both affected by the underlying disorder. Perception of such an underlying cause was the rationale behind the structural physiognomy of Lombroso, who believed that people who offended did so because they were atavistic throwbacks, essentially resembling gorillas (Atwan, 1984).

Unlike self-fulfilling prophecies and expressional physiognomy, accurate perception of an underlying biological cause would not necessarily be linked to developments in appearance or personality over time. Thus cross-sectional studies that link childhood appearance to personality might provide particular indications of underlying biological causation. For instance, an observational study by Langlois and Downs (1979) showed that physically unattractive five-year-olds acted more aggressively than their more attractive peers. Although the cause of this relationship cannot be definitely determined, as they may have been responding to contemporaneous differential treatment and a period of five years could arguably allow the development of a self-fulfilling prophecy, this finding provides some support for a biological cause of both physical attractiveness and aggression. Additional support is given by the finding that minor physical anomalies in new-born infants, including facial anomalies, predict short attention span, aggression and impulsiveness at the age of three (Waldrop, Bell, McLaughlin, & Halverson, 1978).

Only one attempt has been made at determining the accuracy of perceptions from photographs of non-adults. Kessen (1957), using nude, full-body photographs of 10 children aged between 14 to 23 months, found significant associations between the mean rankings of perceived intelligence, activity, mental health and sociability made by stranger judges and the validity criterion of non-physiognomic rankings given by a child psychologist. This psychologist had studied the children intensively over their first two years. The results suggest that people can perceive attributes that are biologically determined through structural physiognomy. Unfortunately however, the portrayal of the children's whole bodies, lack of controls regarding facial expression and the absence of other research into perceptions of this age group all make the results of this study hard to interpret.

Finally, another extremely plausible reason for finding a perceived/real personality correlation is that participants base their judgments on uncontrolled variables in

studies which themselves correlate with the predicted characteristic but which are not traditionally regarded as physiognomic in nature. These could include clothing, grooming, traces left by the environmental like sun-bleached hair, weathered skin, scars, tattoos etc., eye-glasses, race, sex, age, bodily physique, instructions to smile, lighting or even the type of photographic paper used (Alley, 1988; Bull & Rumsey, 1988; Reis et al., 1990; Thornton, 1943). For example, Alley (1988) discusses a study in which students accurately judged peers on their extra-curricular activity levels and athleticism from yearbook photographs and concludes that such "social" judgments ("nerds" or "jocks") were probably based on ascertaining physical attributes such as good looks and bodily stature (judged from facial hints). Because these physical qualities could have partially determined social standing, Alley (1988) argues the degree of accuracy found had little to do with true physiognomic judgments of character/sociability. A similar process could perhaps be behind accuracy in distinguishing muggers from fraudsters (e.g. Thornton, 1939) and might partially explain the results of the replicated sorority studies described by McCabe (1928) (see above). It is extremely conceivable, for instance, that traits such as neatness, refinement, snobbishness and vulgarity might covary with hairstyles and general grooming and this covariance could form the basis of modestly accurate pseudo-physiognomic inferences.

In relation to Zebrowitz's (1997) model of accurate consensus, these uncontrolled variables could operate in many of the links between appearance, psychological traits and the environment. For instance, antisocial psychological traits might lead to indifference regarding grooming, or, alternatively, antisocial psychological traits and poor grooming could be the result of a common environmental cause of association with deviant peers. Poor grooming might even result in unemployment, a different social environment and a subsequent introduction into antisocial thinking. Thus, using Zebrowitz's (1997) terminology, grooming could be seen as an accurate indicator of antisocial traits due to a Dorian Gray effect, a shared environmental cause or a self-fulfilling prophecy. While some of these variables may overlap with stereotypical explanations for judgment consensus, none is relevant to expressional or structural physiognomy, being unrelated to differences in permanent facial features that are caused by idiosyncratic aspects of character or biology.

The above review provides general supportive evidence that character can be accurately judged from static facial appearance and introduces some of the explanations that have been given for that accuracy. A more circumscribed area in which research suggests a potential link between a person's facial appearance and a socially important behavioural outcome is that of forensic psychology. This area provides a locus for the present study. Figure 1 presents a basic model of possible explanations for an empirical correlation between facial appearance and criminality, which could be related to criminal facial stereotypes or accurately perceived through physiognomy. A similar model could be drawn to represent other aspects of character or behaviour.

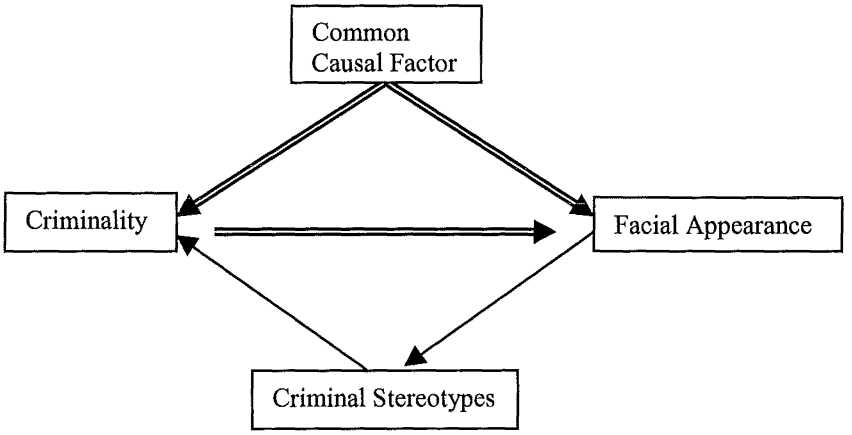


Figure 1. Possible explanations for a relationship between fixed facial appearance and criminality. Double lines indicate pathways associated with expressional or structural physiognomy.

The pathway that is associated with structural physiognomy is that of a common factor causing both criminality and permanent facial appearance. This is typically imagined to be a biological predisposition toward crime (Zebrowitz, 1997). Expressional physiognomy is represented by the pathway of criminality causing facial appearance; a variant of the Dorian Gray effect whereby a criminal lifestyle moulds personality, which in turn changes facial features through habitual muscle tonus. Of

course, both of these pathways could also involve non-physiognomic explanations for a consensus between judges regarding faces and criminality that correlates with the criminal actions of those being judged. For instance, membership in a criminal gang could be a common environmental factor causing both unlawful activity and facial tattoos that are stereotypically associated with criminality. Similarly, a history of physical assault might result in facial scars. Such possibilities would be confounding variables for any research into the accuracy of perceiving the direct expression of underlying personality from the resting face.

The major part of research involving facial appearance and criminal propensity has been concentrated upon the lower, non-physiognomic half of this diamond-shaped model. This includes the effects of appearance stereotypes on findings of guilt within the criminal justice system. The same pathway is also relevant to the concept of the self-fulfilling prophecy, where a person's facial appearance is seen to be stereotypically criminal and this creates a social environment which pushes him or her toward crime. Thus the "Facial Appearance" to "Criminal Stereotypes" to "Criminality" pathway in Figure 1 can be conceptualised as socially constructed, rather than involving any physiognomic perception of an underlying criminogenic personality flaw. Accuracy in facial judgments could be credited to stereotypes which bias the judging participants' attributions of character but which also influence the development of personality in those people whose faces are judged, or even determine whether or not they were convicted and labelled as criminals. In short, stereotypic beliefs could explain any correlation between real and perceived criminality by constituting causes of both the judges' attributions and the targets' crimes.

Connecting Crime, Criminal Stereotypes and Facial Appearance

At present, only some of the connections in Figure 1 have been even partially established. Indeed, the mere existence of a link between facial appearance and criminal propensity that could form the basis of accurate judgments of criminality from static faces is considerably less than certain. However, as for the data relating to the facial perception of traits in general, there is evidence that gives some support to valid physiognomy. Namely, participants can accurately estimate some characteristics that are specifically related to criminality from assessments of fixed

facial features. Certain aspects of actual facial structure and broad facial descriptions like physical attractiveness have also been specifically related to antisocial behaviour. Again, in consonance with the discussion of physiognomy in general, this evidence is also consistent with accurate consensus being driven by the effects of stereotyping upon the development of criminal careers. Nonetheless, it provides reason to suspect that criminal propensity might prove to be perceivable through physiognomy.

Facial appearance and criminal stereotypes. Previous research has consistently revealed that, when given a selection of faces to choose from, participants identify particular faces more than others as belonging to people who have committed various (but not all) specific criminal offences (Bull & Green, 1980; Goldstein et al., 1984; Shoemaker et al., 1973; Yarmey, 1993). This has universally been interpreted to demonstrate that people hold facial stereotypes for certain types of criminal. However, such an interpretation ignores the possibility that the participants accurately perceived differences in personality and related these to their judgments about whether a person was a mugger, a thief, a drug user and so on. Broader facial descriptors such as physical attractiveness (Saladin, Saper, & Breen, 1988) and babyfaceness (Berry & McArthur, 1986; Berry & Zebrowitz-McArthur, 1988a, b) have also been linked to attributions of criminality and criminal intent.

Facial appearance and crime. While, facial stereotypes have been shown to have an effect upon eyewitness testimony (Yarmey, 1982) and to mediate judgments of guilt and innocence beyond the influence of simple physical attractiveness (Macrae & Shepherd, 1989), most research that supports a proximate potential relationship between facial stereotypes and crime has been centred around the concept of physical attractiveness. Within the context of the self-fulfilling prophecy, there is a large body of research into the effects of physical attractiveness on the criminal justice system. A bias toward the physically attractive might help explain any correlation found between crime, facial appearance and stereotypes by implying that facial stereotypes lead to convictions through differential treatment, and also that past convictions place people on the road to further offending.

The method of choice when investigating the effect of defendant physical attractiveness on the mechanisms of justice has been a mock juror procedure modelled

on the seminal study by Landy and Aronson (1969). This procedure, at its simplest, involves providing a photograph of the "defendant" along with a written case summary and asking participants to pass verdict or quantify a sentence. The degree of physical attractiveness is typically determined by a truth-of-consensus method, where an earlier sample of people rates a selection of photographs on attractiveness and manipulation checks are made to ensure that the experimental sample shares their perceptions.

Researchers who have used the mock juror procedure have discovered a leniency effect for physically attractive defendants (for a meta-analytic review of this area see Mazzella & Feingold, 1994). Leventhal and Krate (1977) provide the most convincing experimental example of the physical attractiveness leniency effect. They used a procedure involving multiple photographs and scenarios, which controlled the nature of the crime committed, the race and sex of both the defendant and the participants, plus the economic status of the defendant. They found a significant relationship between the lengths of sentence imposed by the participants and the defendants' attractiveness. Leventhal and Krate (1977) conclude that "how a person looks can influence the term of punishment for an illegal activity as set by members of a mock jury" (p. 317).

While some studies have found a simple main effect for defendant attractiveness, showing no interactions with other independent variables (e.g. DeSantis & Kayson, 1997; Jacobson, 1981), further research has indicated that multiple variables may have an influence on the physical attractiveness leniency effect. These variables include, for example: the crime perpetrated (Sigall & Ostrove, 1975), the sex of the participant (Deitz & Byrnes, 1981; Efran, 1974; Mazzella & Feingold, 1994; Moore, Wuensch, Hedges, & Castellow, 1994; Wuensch, Castellow, & Moore, 1991), the race and culture of both defendant and participant (Wuensch, Chia, Castellow, Chuang, & Cheng, 1993), group deliberation (MacCoun, 1990), and the physical attractiveness of the participant (Darby & Jeffers, 1988). Thus physical attractiveness is not entirely consistent in its effect upon mock jurors, nor has a leniency effect always been found at all (e.g. Boor, 1976; Gerdes, Dammann, & Helig, 1988). Nonetheless, the mock juror research does indicate that facial appearance may have

an influence on responses from the environment of the justice system, which could contribute to the development of a self-fulfilling prophecy.

The mock juror method has been criticised extensively, often centred on three major weaknesses: the use of a student sample, data which is typically obtained on an individual basis rather than from group deliberations, and, the participants knowing that their decisions have no real-life consequences (Bray & Kerr, 1982). There is also concern that the simplistic experiments do not involve enough of the social and legal context of a true court case to reproduce the phenomenon, jury decision making, that is under investigation (Greenwood, 1983) and, more specifically, that many of these experiments use sentencing as a dependent measure when real jurors in a criminal trial pass a verdict on guilt (Bray & Kerr, 1982). Nonetheless, as argued by Mook (1983), such experiments can indicate a possible *process* without requiring the strict adherence to representative sampling needed in, for instance, survey research, where the purpose is to determine the actual characteristics (i.e. verdicts) of the population. The process revealed by the mock juror experiments is a biasing effect of defendant attractiveness on the judgments of participants, even although participants report a general belief that defendant physical attractiveness should be irrelevant to jury decisions (Efran, 1974). It follows from Mook's (1983) position that, while it would be improper to translate experimental effect sizes into predictions of actual jury decisions, it might be concluded that physical attractiveness does affect jurors to some extent and that certain other factors, found experimentally, could interact with that influence.

In the only two field-studies investigating the attractiveness leniency effect (Stewart, 1980, 1985) observers rated real defendants on physical attractiveness during their criminal trials. The physical attractiveness ratings were then compared to the severity of punishments received and the direction of verdicts. In the first study (Stewart, 1980), 74 defendants were observed over two years at various courts in the United States. The relationship between verdict and physical attractiveness was not significant, suggesting that the juries were impartial. Judicial sentencing was another matter entirely. The length of sentences significantly negatively correlated with physical attractiveness and this remained after statistically controlling for seriousness of the offence. In the 1985 study (of 60 defendants) the single rating of physical

attractiveness was elaborated into physical attractiveness, cleanliness, neatness and quality of dress. Again, even after the seriousness of the crime committed was statistically controlled, there was a significant correlation between a combined score for these factors and sentencing but not for verdict.

Despite the lack of effect on verdicts, the effect of attractiveness on sentencing in the observational studies lends support to the external validity of the physical attractiveness leniency effect found in mock juror experiments. These studies do, after all, tend to use sentence length as the main dependent variable because dichotomous verdict may be too statistically insensitive. Fortunately, however, it appears that the leniency effect is too small to be manifested in real-life jury decisions. If the attractiveness halo does cause some instances of crime through leading to convictions for those who are unattractive and influencing other interactions within the justice system that might lead to self-fulfilling prophecies, the relative importance of the effect is probably fairly low. The field studies also suggest that judges and other legal officers are more likely than jury members to be conduits for this improper influence.

To this quite limited evidence, supporting a connection between facial appearance and crime through the auspices of stereotypes held within the criminal justice system, can be added some data pertaining to physical appearance and actual behaviour. As already mentioned, physically unattractive five-year-old children have been observed to act more aggressively than their more attractive peers (Langlois & Downs, 1979). This is consistent with a causally ambiguous correlation between later criminal aggression and facial appearance, which fits the stereotype that criminals are physically unattractive. Meanwhile, some reduction in recidivism appears to follow surgery performed to correct facial deformities in prison inmates (see reviews by Bull, 1982; Bull & Rumsey, 1988; Thompson, 1990), which highlights the potential self-fulfilling effects of facial abnormality.

A more provocative finding is that membership in four theoretical categories of juvenile delinquency (inadequate-immature, neurotic-disturbed, unsocialised-psychopathic, and socialised-subcultural) has been associated with significant differences in facial attractiveness (Cavior & Howard, 1973). As well as showing

differences between themselves that coincided with theoretical divisions, the juvenile delinquents Cavior and Howard (1973) investigated, when taken together, were rated as significantly less physically attractive than a control group of high school students. The same connection between delinquency and low attractiveness is confirmed to some extent by another study. Agnew (1984) compared self-reported delinquency behaviour between those who were rated as having the worst and best general appearances from a larger sample of 11th grade males participating in a longitudinal survey. The children with poorer appearances rated themselves as significantly more delinquent. However, the rating of appearance was performed after a two hour interview and involved more than simple facial attractiveness (including dress and grooming), therefore the relevance of Agnew's (1984) results to imputing a link between crime and facial appearance is somewhat diminished (Bull & Rumsey, 1988).

Likewise, in Stewart's two field-studies (1980, 1985, mentioned above) in which observers rated real defendants on attractiveness during their criminal trials, an ancillary finding in both investigations was that there was a significant negative correlation between defendant attractiveness and the seriousness of the crime committed (i.e. murder cf. rape cf. indecent assault etc.). This implies a connection between facial appearance and aspects of actual criminal offending. However, the observers were exposed to more than simply the defendants' faces at rest, so it is uncertain exactly what the ratings of attractiveness were based upon, and the possibility remains that "cognitive biases make it consistent to perceive a violent criminal as unattractive" (Stewart, 1985, p. 377).

In contrast, some data relating both physical attractiveness and babyfaceness to criminal behaviour actually contradict the stereotypical beliefs of the attractiveness halo and the warmth and submissiveness ascribed to those with baby-like faces. Zebrowitz, Andreoletti, et al. (1998) and Zebrowitz and Lee (1999) reanalysed information from a longitudinal crime causation study which compared a sample of juvenile delinquents to a matched sample from proximate public schools at the ages of 0-17, 17-25 and 25-32 years. In a cross-sectional analysis of boys between the ages of 10 to 17 years, physical attractiveness and babyfaceness, as rated by students from facial photographs, were found to have small effects upon social relationships that had

been related to delinquency for the same sample (Zebrowitz & Lee, 1999). These social relationships were maternal supervision, appropriate paternal discipline, parent-child attachment, peer relations and gang membership. Among the delinquents, those who were more babyfaced or attractive tended to suffer poorer relationships with their families, interpreted as a contrast effect arising from broken expectations. This was contrary to predictions based on attractiveness and babyfaceness stereotypes. More babyfaced and attractive non-delinquents had better family relationships than their less attractive, more mature-faced, non-delinquent peers.

In further contradiction of stereotypical expectations, Zebrowitz, Andreoletti et al. (1998), conducting an analysis involving all three age groups, found that low socioeconomic status babyfaced boys were more likely than their mature-faced peers to be delinquent. Also, between the ages of 17 and 25 years, the more babyfaced delinquents faced criminal charges more often than mature-faced delinquents. Meanwhile, attractiveness was positively related to delinquency (contrary to the halo effect), but only for boys who were either of higher socioeconomic status or who were more muscular, while between 17 and 25 the less attractive committed more crimes, but only if they were also highly muscular. Disregarding interactions, babyfaceness and physical attractiveness were unrelated to being charged with an offence up to the age of 17 or after 25, and, for attractiveness alone, also between 17 and 25. As Zebrowitz and Lee (1999) surmise, "It appears that there may be a downward spiral when baby faced adolescents violate the well-documented expectency that they will be submissive, warm, and weak. Their violation of these benign expectancies engenders social relationships that, paradoxically, may make subsequent antisocial behavior more frequent for the baby faced than the mature faced" (p. 583). In the context of discussing accuracy to be found in criminal stereotypes, data from this sample actually largely contradict stereotypes regarding babyfaceness and attractiveness but they do support a connection between facial appearance and criminal propensity that could potentially be perceived.

Very little research has been conducted that does not specifically relate to those global facial descriptions, that is physical attractiveness and babyfaceness, that have been implicated in facial stereotypes. However, one line of enquiry, involving the effects of minor physical anomalies suggests a biological deficit affecting both criminal

propensity and facial appearance which has particular relevance to structural physiognomy. These anomalies, while not all facial, mainly include aspects of facial structure such as widely spaced eyes and asymmetrical ears. Minor physical anomalies have been associated with hyperactive and problem behaviour in non-clinical children (Firestone & Peters, 1983; Halverson & Victor, 1976; Waldrop et al., 1978; Waldrop, Pedersen, & Bell, 1968), with increased aggression in male college students (Paulhus & Martin, 1986) and, most strikingly, with recidivism relating to adult violent criminal behaviour (Kandel, Brennan, Mednick, & Michelson, 1989). A common biological cause is probably the most provocative potential reason for a connection between crime and facial appearance that could lead to accurate consensus in facial judgments. Although there is a body of research supporting a biosocial approach to antisocial behaviour, where biological deficits interact with the social environment to produce crime (Raine, Brennan, & Farrington, 1997), apart from the connection made between minor physical anomalies and criminal conduct there seems to be no other research specifically directed at linking criminal acts and facial appearance to a shared biological cause. It also remains to be determined whether people actually advert to minor physical anomalies of the face and use them as indications of character or behaviour.

In summation there is some evidence suggesting a connection between facial appearance and crime that could form the basis for accurate perception. Facial appearance has been weakly linked to crime by research indicating that jurors may convict the less attractive more frequently. This could possibly be extrapolated to a general bias in the justice system that encourages the development of self-fulfilling prophecies but such an extension would have little foundation. There are also a small number of studies that have linked physical attractiveness to actual juvenile delinquency (Cavior and Howard, 1973; Agnew, 1984) and the seriousness of adult offending (Stewart, 1980, 1985). Others, while finding a slight connection between appearance and criminal behaviour, and social relationships linked to delinquency, produced results that contradicted attractiveness and babyfaceness stereotypes (Zebrowitz, Andreoletti et al., 1998; Zebrowitz & Lee, 1999). There is also some evidence that specific facial features, described as minor physical anomalies, are related to problematic behaviour and recidivism. This, in particular, supports the

possibility of a biological cause allowing for valid structural physiognomy relating to criminal propensity.

Accuracy in Perceiving Aspects of Criminality from the Face at Rest

Whereas there is some evidence of an actual relationship between facial appearance and crime, support for the accurate facial perception of criminality is sparse. Only two studies have displayed a degree of accuracy in facial discriminations between people who have committed different offences (Kozeny, 1962 cited by Bull & Green, 1980; Thornton, 1939). It should be emphasised from the outset that neither of these studies compared the perception of criminal faces to those of non-criminal controls and thus they associated faces with particular crimes, not with criminal propensity *per se*.

Bull and Green (1980) refer briefly to an earlier study, published in German, in which Kozeny (1962) produced composite portraits representing different categories of criminals. These were based upon photographs of 730 convicted criminals and the types of crimes they had committed. Kozeny (1962) found that the physiognomic characters of these composite portraits differed significantly, and Bull and Green (1980) take these results to suggest that there is "some relationship between physical appearance and criminality" (p. 80). However, their truncated descriptions of Kozeny's (1962) procedures and results lack sufficient detail to allow a critical assessment of this conclusion.

Thornton (1939), in the only study published in English, selected 20 people out of a number of case records from a state penitentiary. No reference was made to any photographs. The selected people had no apparent prison record apart from a single conviction, were white males, and collectively represented a variety of crimes. Following selection, the photographs that were taken at their incarcerations were retrieved, converted to head-only facial stimuli and presented to 175 student participants. Their task was to match the correct crime to each face, from a list of four crimes. Each of ten lists was used twice, paired with two different photographs in such a way that the correct item for one photograph served as a control item for the other and *vice versa*. A comparison of the frequencies at which correct or control

items were chosen showed that students could pick the correct crime more often than expected by chance.

Thornton (1939) took efforts to choose his faces without looking at any photographs, in order to prevent his own stereotypes from biasing the selection. He rightly feared that if photographs were viewed during selection the study would degenerate into a simple determination of the similarity between his and the participants' stereotypes. Unfortunately, this meant that possible confounding variables such as age, wearing glasses, grooming, environmental traces and so on (Alley, 1988; Bull & Rumsey, 1988; Thornton, 1943; Reis et al., 1990), as well as variables associated with facial stereotypes, could have been the factors which cued accurate choices. Thus the accuracy shown in the choices of his participants is ambiguous in relation to expressional or structural physiognomy. Thornton (1939) notes that only one target was wearing glasses and no clothing or ornaments were apparent in the photographs but raises the possibility of age as a confounding factor and does not discuss any other features of the facial stimuli. It is also quite possible that, despite being imprisoned only once, the targets had in fact committed more than one crime each, calling into question the validity of the criterion used to judge a correct match.

In addition to these two studies relating to the perpetration of actual offences there is some mixed evidence relating to accuracy in perceiving honesty, a characteristic that is intrinsic, through its absence, to much criminal behaviour. A longitudinal investigation by Zebrowitz et al. (1996) indicates a very limited degree of accuracy in facial perceptions of honesty, one forming a pattern that is more consistent with accuracy due to the effects of stereotyping than due to physiognomy. Zebrowitz et al. (1996) examined data from three prospective longitudinal studies that provided information, including photographs, about a middle class, Caucasian sample from childhood to 50-60 years. Ratings of honesty were made from the photographs by students and were compared to a real honesty criterion consisting of archival Q-sort ratings made by clinicians ("is guileful and deceitful" and "appears straightforward, forthright, candid in dealings with others"). The Q-sort ratings had been based on interview, questionnaire and observational data. The researchers found that there was a general lack of correspondence between real and perceived honesty, indicating a lack of accuracy in facial judgments. An exception was found in marginally

significant negative relationships for those assessments made at puberty, showing that the facial judgments were *inaccurate*. In addition, women who were stable in real honesty exhibited an artifice effect. The facial judges were deceived into rating those women who were actually less honest in early life as more honest later on, probably through deliberate manipulation of facial appearance by the use of make-up and other grooming aids. The only accurate facial judgments were made of a *post hoc* sub-grouping of adult men. Facial judgments of adult men displayed a marginally significant positive relationship to real honesty, provided they had remained stable in perceived honesty between their early years (averaged across childhood, puberty and adolescence) and later years (averaged across 30s and 50s). Men who showed a stable appearance of honesty in their younger years also exhibited a marginally significant self-fulfilling prophecy effect. That is, men who looked more honest in their early years became more honest later. Thus the results were consistent with what little accuracy that was found in perceptions of honesty having resulted from a self-fulfilling prophecy effect.

Such equivocal results regarding accuracy can be contrasted to those of Bond et al. (1994), who found that strangers' ratings of dishonesty based on facial photographs were positively correlated with a willingness to participate in a number of experiments involving deception. This occurred even although the required deception was highlighted by the experimenters and other choices were made readily available. Berry (1990) also reports a study in which strangers' impressions of honesty from facial photographs correlated with ratings of the targets' honesty by classmates of nine weeks' standing. However, some caution is warranted in interpreting this outcome, as the students' level of acquaintance may not have been very high after only nine weeks, challenging the validity of the real honesty criterion. It seems plausible that the classmates also based their judgments of honesty partly on facial impressions, given the limited personal knowledge they had of their fellows. If this was the case then the supposed accuracy might be entirely attributable to facial stereotypes for honesty bringing about consensus between stranger and classmate ratings without the involvement of any kernel of truth. Overall, these three studies do not provide particularly strong support for accuracy in facial perceptions of honesty, and the specific situation in which accuracy was shown in the study by Zebrowitz et al.

(1996) was most consistent with a stereotypical explanation of consensus rather than expressional or structural physiognomy.

Psychopathy: A Potential Link between Facial Appearance and Crime that could be perceived through Physiognomy

Of the possible candidates for a common cause of both crime and facial appearance that can be gleaned from the current research literature the most promising is probably the personality disorder of psychopathy. The personality traits and behaviours related to psychopathy might prove to be perceivable through structural or expressional physiognomy. There are six reasons for maintaining optimism regarding the potential of psychopathy as an area for investigating physiognomy in relation to crime. They are: the availability of a reliable means of identification; a connection to criminal behaviour; indications of a possible biological cause; an argued association with hyperactivity in children, which has, in turn, been linked with minor physical anomalies; the presence of extreme personality traits; and, a similarity to Machiavellianism, which has been shown to be accurately judged from photographs.

First, psychopathy can be reliably identified. The Hare Psychopathy Checklist-Revised (PCL-R) (Hare, 1991) is a clinician rating scale which shows both reliability (Alterman, Cacciola, & Rutherford, 1993; Hare et al., 1990) and validity (Forth, Hart, & Hare, 1990; Hart & Hare, 1989; Serin, Peters, & Barbaree, 1990) in assessing psychopathy. This enables a progressive accumulation of comparable research dealing with the same, well-defined, group of people.

Second, scores on the PCL-R predict various aspects of criminal behaviour such as the onset, frequency, severity and types of offending, as well as recidivism (e.g. Cornell et al., 1996; Forth et al., 1990; Hart, 1998; Hart, Kropp, & Hare, 1988; Hemphill, Hare, & Wong, 1998; Hemphill, Templeman, Wong, & Hare, 1998; Serin, 1991, 1992; Serin et al., 1990). Hence, psychopathy has already been linked to crime. Most particularly, both extreme criminality and psychopathy find a strong juncture in the persons of sexually motivated serial killers, who also possess enough notoriety to have their photographs published but not so much fame that there is wide exposure of their images to the public of New Zealand. Sexual serial killers appear to validate the

popular image of the Caucasian male with a psychopathic and sadistic personality who chooses to act out his sexual fantasies for the purpose of his own pleasure (Fox & Levin, 1994; Gekoski, 1998; Levin & Fox, 1986; Warren, Hazelwood, & Dietz, 1996). Empirical support for this widely held view, that all sexual serial killers are psychopathic, can be found in a study by Geberth and Turco (1997). After identifying 68 male serial killers who violated their victims sexually from a search of newswire services in the United States, Geberth and Turco (1997) found that every one of them fulfilled DSM-IV (American Psychiatric Association, 1994) criteria for both sexual sadism and antisocial personality disorder (APD). Although APD is not equivalent to psychopathy as determined by the PCL-R (Hare, 1996), it was originally intended to operationalise the classical (e.g. Cleckley, 1964; McCord & McCord, 1956) conceptions of the disorder. Those with psychopathy do appear to form a subset of those diagnosed with APD (Lilienfeld, 1994) and thus, while Geberth and Turco's (1997) study does not directly find that sexual serial killers are all psychopathic, it certainly does support this view.

Third, the group identified by use of the PCL-R has been shown to differ from comparisons across many dimensions (Hart & Hare, 1997). Those labelled psychopathic have, for example, been found to: have abnormal emotional responses (e.g. Hare, 1998; Patrick, Bradley, & Lang, 1993; Patrick, Cuthbert, & Lang, 1994; Patrick, Zempolich, & Levenston, 1997); evidence less lateralised linguistic processing (e.g. Raine, O'Brien, Smiley, Scerbo, & Chang, 1990); show unusual reward dominance, passive avoidance learning, or response modulation (Harpur & Hare, 1990; Howland, Kosson, Patterson, & Newman, 1993; Kosson & Newman, 1986; Newman, 1998; Newman & Kosson, 1986; Newman, Patterson, & Kosson, 1987; Newman, Schmitt, & Voss, 1997; Scerbo et al., 1990); and, possibly exhibit subtle neurobiological problems (see Dolan, 1994; LaPierre, Braun, & Hodgins, 1995 cf. Hart, Forth, & Hare, 1990). These findings, plus evidence that psychopathy represents a discrete class, or taxon, rather than forming part of a continuous personality dimension (Harris, Rice, & Quinsey, 1994), serve to argue for the potential presence of a biological abnormality. Such an abnormality might also affect facial appearance and thereby allow structural physiognomy.

The fourth reason advocating psychopathy as a worthy area for investigating physiognomy in relation to crime is an hypothesised association between adult psychopathy and childhood symptoms of attention deficit hyperactivity disorder. Lynam (1996) proposes that children who manifest symptoms of hyperactivity, impulsivity and attention problems (HIA) in combination with conduct problems are not only at risk for chronic offending (e.g. as suggested by Moffitt's, 1993, dual taxonomy of adolescent-limited and life-course-persistent antisocial behaviour) but can be described properly as fledgling psychopathic offenders. He supports this contention by reviewing studies comparing such children with others who have either conduct problems or HIA or neither. He presents evidence that the "fledgling psychopaths" display more antisocial behaviour as adolescents and receive more diagnoses of APD as adults, that their relatives are more antisocial, that their patterns of antisocial behaviour (onset, frequency, severity) are those that have been identified as risk factors for adult APD, and, most tellingly, that they share the same sorts of deficits as adult male psychopaths in laboratory, psychophysiological and performance measures. . If Lynam's (1996) argument is correct, and this is placed within the context of the further tie between hyperactive symptoms and minor physical anomalies, it supports a link between psychopathy, and biological causes of facial appearance. This link, in turn, lends credence to an expectation that psychopathy might be subject to structural physiognomic identification.

Fifth, according to its manual (Hare, 1991) the PCL-R aims to measure a personality-based view of psychopathy. Although, it has been argued that, to some extent, the PCL-R actually confuses personality-based and behaviour-based definitions of psychopathy (e.g. Lilienfeld; 1994, 1998) it clearly does involve a group of extreme personality traits. Factor analyses have indicated two correlated but distinct factors in responses to the PCL-R (Cooke & Michie, 1997; Hare et al., 1990) and its predecessor the PCL (Harpur, Hare, & Hakstian, 1989). The first of these factors is said to describe personality, including: superficial charm, a grandiose sense of self-worth, pathological lying, manipulativeness, lack of remorse or guilt, shallow affect, lack of empathy, and failure to accept responsibility (Hare et al., 1990). It is found more frequently in people showing higher levels of psychopathy than the second factor, a chronically unstable and antisocial lifestyle (Cooke & Michie, 1997). These "egocentric, manipulative, and callous traits" also exhibit more cross-sectional

stability across the life-span and seem particularly fundamental to psychopathy (Harpur & Hare, 1994, p. 604). Thus, even if psychopathy is not based on a biological abnormality that is perceivable through structural physiognomy, these traits could be expected to contribute to the traces of habitual facial expression upon which expressional physiognomy is said to depend. Assessment of accuracy in rating the traits that make up the first factor of the PCL-R provides an obvious opportunity to test the physiognomic perception of personality.

Finally, Cherulnik et al. (1981) determined that students have an ability to accurately judge levels of Machiavellianism from photographs. McHoskey, Worzel and Szyarto (1998) argue that Machiavellianism and psychopathy are actually equivalent constructs developed in parallel by two different branches of psychology, social and clinical. They refer to conceptual similarities between psychopathy and Machiavellianism, that is, shared core qualities of affective detachment, intact reality contact and manipulativeness. They also point to past research showing associations between Machiavellianism and all of the traits that form the PCL-R, and complement this with their own demonstration of a positive association between measures of Machiavellianism and of primary and secondary psychopathy in a university sample. If McHoskey et al.'s (1998) argument is correct, then this means that there is already some evidence that psychopathy is a perceivable common factor behind both criminal propensity and facial appearance. This possibility alone provides a good reason to identify psychopathy as a worthwhile area for future investigations into the validity of physiognomic perceptions of criminality.

Summary

Contrary to the strongly held beliefs of past research psychologists, there is a body of evidence which supports the accuracy of judgments made regarding the characters of people from their fixed facial features. An ecological perspective on person perception predicts that people should be able to directly perceive affordances, that is opportunities for adaptive behaviour offered by others, and provides a theoretical basis for physiognomy. Cognitive theories of consensus in facial judgments highlight the importance of over-generalisation from the adaptive perception of variables such as age, fitness, emotional expression and gender. It is unclear whether accuracy is

due to physiognomy or the self-fulfilling effects of stereotypes. The evidence relating to accuracy in facial judgments does not directly speak to whether that accuracy is due to expressional physiognomy, structural physiognomy, or over-generalisation, although the self-fulfilling prophecy effect does have the greatest volume of supporting literature.

More specifically, there is evidence relating facial appearance to crime. As for the general literature concerning facial judgments of underlying traits, judges share a degree of consensus, in this case regarding which faces are perceived to be those of people who have committed criminal acts. This consensus has largely been interpreted as resulting from people holding stereotypes regarding the faces of criminals. An alternative interpretation suggests that judgment consensus is at least partly due to real relationships between facial appearance and crime. Evidence supporting proximate links between facial appearance and crime is largely restricted to experimental designs investigating the effects of a broad facial description, physical attractiveness, upon the deliberations of juries. This research has revealed a complex influence of defendant physical attractiveness on juridical decisions, one that interacts with many other variables and which has limited external validity in predicting real-life outcomes in the criminal justice system. Such a link does not provide confidence in the actual development of self-fulfilling prophecies. However, a few studies have found associations between antisocial behaviour and physical attractiveness, babyfaceness and minor physical anomalies, some of which indirectly support accuracy in judgments of criminality from permanent features of the face.

Evidence that aspects of criminal behaviour can be accurately perceived from the face at rest is limited to two studies involving actual offenders and some analogue studies concerning dishonesty. There is no research that has investigated the accuracy of physiognomic impressions of general criminal propensity, although one study did confirm the accurate perception of Machiavellianism which, if viewed as a parallel concept to psychopathy, may be related to criminal behaviour. Again, there is no evidence that distinguishes between the potential causes of accurate facial judgment, although the results of one study were consistent with accuracy in facial perceptions of honesty having resulted from the effects of a self-fulfilling prophecy. Psychopathy, as defined by the PCL-R holds promise as an area of investigation into the accurate

perception of criminal propensity, and appears peculiarly conducive to the explanations of structural or expressional physiognomy.

The Rationale for this Study

The aim of the present study was to find evidence supporting the direct physiognomic perception of criminal propensity and personality traits. An ecological approach to person perception predicts that criminal propensity and psychopathy are directly perceivable through physiognomy because such perception would afford adaptive action in others. Furthermore, the results of two studies using convicted criminals as targets for facial judgments and evidence supporting accuracy in facial perceptions of Machiavellianism, also indicate that criminality and the personality disorder of psychopathy provide good opportunities to test the validity of physiognomy.

It was reasoned that differential support for the physiognomic explanations of accuracy in facial judgments could be obtained by ruling out the effect of those variables related to major facial stereotypes, and through these, self-fulfilling prophecies. Control could also be placed upon other non-physiognomic confounds that might explain consensus among ratings on psychopathic traits and criminality, without recourse to physiognomy. These non-physiognomic physical qualities of the photographs or their subjects might have been the reasons for the accurate discriminations found by Thornton (1939) and many other researchers. Physical qualities that are not relevant to physiognomy include relatively mundane aspects of the images themselves, such as facial orientation, picture quality and lighting, in addition to certain aspects of the target people. These include standards of personal grooming plus, most notably, the stable personal attributes that are identified by Zebrowitz (1997) as potential focuses of the processes of over-generalisation: age and babyfaceness, sex, ethnicity/race, identity, emotion and physical attractiveness (indicative of fitness). While any research design that is reliant upon identifying and excluding other potential explanations cannot prove that there is a physiognomic basis for judgment accuracy, it can potentially help to indicate whether further research into physiognomy would be worthwhile. By ruling out other explanations of accuracy in facial judgments, the feasibility of physiognomy is indirectly supported.

It was predicted that participants asked to make judgments from facial photographs would rate sexual serial killers as having greater propensity toward crime and possessing more psychopathic personalities than *prima facie* non-criminals. Facial photographs of sexual serial killers were used on the assumption that they exhibited higher levels of criminal propensity and psychopathic personality traits as a group than the *prima facie* non-criminals. This assumption of group differences did not rely upon all of the killers technically fulfilling the requirements of a diagnosis of psychopathy. A comparison group of *prima facie* non-criminals was thought necessary because perception of general criminal propensity was the intended object of study, not simply accuracy in distinctions between the specific types of criminal acts performed. The use of sexual serial killers might also be seen to maximise any Dorian Gray effect that crime may have on the face; their crimes and personalities are extreme by any definition.

To summarise, the present study aimed to find support for physiognomy through controlling competing explanations for accuracy in facial judgments. A comparison of facial perceptions of sexual serial killers and *prima facie* non-criminals was selected to test physiognomy because both the ecological approach to person perception and the empirical literature suggest that criminal propensity and psychopathy should be amenable to direct physiognomic perception. It was predicted that, after controlling non-physiognomic differences between the groups, ratings of criminal propensity and the possession of psychopathic personality traits would still be higher for sexual serial killers than for *prima facie* non-criminals. This would support the validity of physiognomic explanations for accuracy in facial judgments.

Method

Participants

The participants were 70 first-year psychology students recruited over a period of seven days from nine laboratory classes, which were part of an introductory course in psychology at the University of Canterbury, New Zealand. Recruitment comprised of a short speech given by the experimenter followed by an invitation to participate in an adjoining room at the termination of the laboratory class. Participation was described as involving the making of multiple ratings from facial photographs, requiring a period of up to an hour for completion. An inducement of $5.00 cash and Instant Kiwi Scratch Tickets to the value of $4.00 was offered for volunteering. Four participants were omitted from the analyses because they correctly recognised some of the target photographs, leaving a final sample of 33 females and 33 males aged from 17 to 53 years (M=21.55, SD=7.07).

Stimulus Materials

A preliminary study was run to select photographic stimuli for the main study. The participants in the preliminary study were 10 post-graduate psychology students (five male and five female) recruited on a single day from a seminar and a computer laboratory in the Psychology Department of the University of Canterbury, New Zealand. Each was a volunteer who was paid $10.00. Their ages ranged from 22 to 42 years of age (M=33.7, SD=6.38).

Classification of a person as a sexual serial killer (SK) was determined by the experimenter from biographical accounts. Classification was dependent upon the fulfilment of two criteria: (1) that the person had committed two or more murders in separate events, and (2) that there was an indication of sexual motivation in at least one of those murders from circumstances at the crime scene. The criteria were themselves based upon the definition of serial murder proffered by the National

Institute of Justice in the United States (NIJ) combined with the method of inferring sexual motivation from observations of the crime scene, as modelled by Ressler, Burgess, Hartman, Douglas and McCormack (1986). The NIJ definition of serial murder is "a series of two or more murders committed as separate events, usually, but not always, by one offender acting alone. The crimes may occur over a period of time ranging from hours to years. Quite often the motive is psychological, and the offender's behavior and the physical evidence observed at the crime scenes will reflect sadistic, sexual overtones" (Brooks, Devine, Green, Hart, & Moore, 1988, cited by Geberth & Turco, 1997, p. 49). Observations at the crime scene indicating sexual motivation include, but are not limited to: victim attire or lack of attire, exposure of sexual parts of the victim's body, sexual positioning of the body, insertion of foreign objects into body cavities, or evidence of sexual intercourse (Ressler et al., 1986).

Facial photographs of 19 adult, Caucasian male SKs with no facial hair or head adornments were found in a search through published materials relating to serial murder (see Appendix 1 for a list of the names of the SKs and the sources from which the photographs were obtained, as well as short descriptions of the historical data that fulfilled the two classificatory criteria given above). From these, 10 SKs were randomly selected and designated as experimental stimuli. Where there was more than one portrait of a SK, the photograph that showed the clearest, fullest representation of his face was chosen. No other characteristics of the photographs (i.e. the features or emotional expressions of the SKs) were considered in the act of selection.

Facial photographs of 179 *prima facie* non-criminals were obtained from the same published materials that were searched for SK photographs and also from books of photographic portraits (see Appendix 2 for a list of the names and occupations of the people in all photographs, when known, and the sources from which they were taken). For consistency, throughout this thesis the photographs are labelled according to their order of listing in Appendix 2, with SK photographs denoted by written words and non-criminals by numerals. All the facial photographs in a source were utilised, provided they were of Caucasian, adult males with no facial hair or head adornments, were not explicitly described as depicting criminals or relatives of SKs, showed

people with their eyes open, and, were neither obscured by another object (e.g. a hand over the face) nor of such poor quality that facial features could not be distinguished. Certain faces were also excluded because they were considered to be very widely recognisable (for instance, former U.S. presidents and members of the British royal family).

The 10 SK and 179 non-criminal stimuli were rated by the experimenter on the following dimensions: age (20s, 30s, 40s, 50+), grooming (poorly to well-groomed on a 5-point scale), picture lighting (dark to light on a 5-point scale), picture quality (poor to good on a 5-point scale), the presence of environmental traces such as scars and tattoos (absent, present, abundant) and facial orientation (left, forward, right; where orientation was judged according to the vector following the nose, and any direction within a 30° angle either side of facing straight ahead was taken as a forward orientation).

Each SK photograph was matched with 10 non-criminal photographs that received the same ratings on age, facial orientation and the presence of environmental traces, and which were rated within one point on the scales for grooming, picture lighting and picture quality. Where there were more than 10 matches the excess were randomly discarded and, as 10 matches for every SK photograph could not be found from the original 179 stimuli, 29 additional non-criminal facial stimuli were collected from further portraiture anthologies and some magazines. These photographs were selected specifically to match those of the SK stimuli with less than 10 matches (photographs ONE, THREE, SIX and SEVEN) on the ratings listed above (see Appendix 2 for names and occupations, if known, and the sources from which they were taken; see Appendix 3 for the ratings given to all 218 photographs). In the case of photograph ONE, fashion magazines were used instead of photographic anthologies because they were expected to be richer sources of average to poorly groomed males in their thirties.

The final 110 photographs were digitally photocopied to a standard size of 4cm across (measured between the furthest two points of uncovered skin in parallel to an imaginary line drawn between the eyes). They were then cropped around the outline

of the head and just below the chin to remove clothing and background cues. Each group of 11 matching pictures was placed on one page of the 10-page preliminary photographic array (see Appendix 4). Within each page the SK image was randomly allotted one of the 11 possible positions (without duplication across the pages) and the 10 matching non-criminal photographs randomly filled the remaining locations. The order of pages was sequentially varied so that 10 different arrays were produced with no particular page following another more than once.

The preliminary questionnaire consisted of two parts (see Appendix 5 for a copy of a preliminary questionnaire). The first section required the participants to disclose their own sex and then identify, name and describe the occupation of any people who they recognised from the preliminary photographic array. The second part involved rating all 110 photographs separately on 1-7 Likert-type scales for physical attractiveness (PA; 1- "very physically unattractive" to 7- "very physically attractive"), babyfaceness (BAB, 1- "very baby-like face" to 7- "very mature face"), and facial expression (EXP; 1- "very negative" to 7- "very positive"). The order of questions in the second part was sequentially varied, with four of the six possible permutations being duplicated over the sample of 10 participants.

The participants were provided with a preliminary photographic array and preliminary questionnaire to take away and complete at their convenience (these were all returned in person within a period of two weeks). The participants were asked verbally to write their age, in years, on the questionnaire.

The photographs that comprised the main photographic array were chosen according to the results of the preliminary study (see Appendix 6). As the person in picture ONE (Jeffrey Dahmer) was identified by two of the 10 preliminary participants, this photograph and its ten matching photographs were excluded from the main study. The non-criminal photographs #55 (Roman Polanski, husband of victim/director), #84 (Edward R. Murrow, radio and television reporter) and #98 (Mikhail Baryshnikov, ballet dancer) were also excluded due to the people portrayed being recognised in the preliminary study. Picture FIVE (Fred West) was described as a "British mass murderer" by one of the two participants who recognised Jeffrey Dahmer, but no name was recalled. It was decided to retain picture FIVE as the risk of recognition by

participants (and hence their exclusion from data analysis) was deemed to be acceptable when weighed against further reducing the number of photographic stimuli in the main study.

Within each cluster of 11 matching photographs, the photograph most similar to the SK was selected by the following procedure. Scores for each attribute (physical attractiveness, babyfaceness and expression) were averaged across participants for every photograph. For every non-criminal photograph the absolute difference from the SK's mean score for an attribute was calculated separately for PA, BAB and EXP. Those three absolute differences in mean scores were then averaged to give a between-subjects indication of perceived similarity to the SK stimulus. This measure of similarity was triangulated against a more individually based analysis, where the absolute difference between an individual participant's ratings of the SK photograph and each non-criminal photograph was calculated for each of the three attributes. Then these scores were averaged across both the three attributes and the participants. In five out of nine cases the same non-criminal photograph was selected by both analyses. For the others, the total mean differences score from each procedure was given equal weighting and the photograph with the lowest combined score chosen. In every situation, the picture finally chosen was ranked as either the most or second most similar to the SK picture by both the between-subjects or within-subjects methods of determining similarity. Over the nine selected non-criminal photographs, the mean between-subjects difference score was 0.36 (*SD*=0.17) and mean within-subjects difference score was 0.79 (*SD*=0.19).

The final 18 photographs were: TWO (Albert De Salvo, SK), THREE (Charles Hatcher, SK), FOUR (Neville Heath, SK), FIVE (Fred West, SK), SIX (Angelo Buono, SK), SEVEN (Theodore Bundy, SK), EIGHT (Ed Gein, SK), NINE (Harvey Carignan, SK), TEN (Ian Brady, SK), #25 (Ludovic Kennedy, unspecified occupation), #115 (Odgen Pleissner, artist), #140 (Ronald Ziegler, President Nixon's press secretary), #153 (Carl Toms, theatre designer), #155 (Alfred Kazin, unspecified occupation), #162 (Henry Moore, sculptor), #172 (Anonymous, wrestler in a gym in Kiev), #176 (Andy Warhol, artist), and #190 (Anonymous, industrial engineer). Comparisons between the SK photographs and the non-criminal photographs selected as most similar to them showed group mean attribute scores to be identical for

physical attractiveness (SK M=2.91, SD=0.73; non-criminal M=2.91, SD=0.60) and non-significantly different for babyfaceness (SK M=4.83, SD=0.78; non-criminal M=4.60, SD=0.86) and facial expression (SK M=3.12, SD=1.31; non-criminal M=3.18, SD=0.81).

The photographs selected from the preliminary study were digitally scanned from their original sources at a resolution of 600dpi and manipulated using Adobe Photoshop 4.0 to be of a standard size (see above) and similar level of darkness (as judged by the experimenter). Picture quality was also enhanced for certain photographs through use of the "blur" and "sharpen" functions, but no facial features were changed. Once more, the photographs were cropped immediately below the chin and around the outline of the head, and then placed into two-page photographic arrays with nine photographs on each page (see Appendix 7 for a copy main study photographic array A). Four arrays were produced (Arrays A, B, C and D), each only differing from the others in the randomly determined order of their photographs.

The questionnaire used in the main study had three parts (see Appendix 8 for a copy of the main study questionnaire, Type I). Four versions of the main study questionnaire were produced (Types I, II, III and IV) which differed only in the order of items. The order of sections remained consistent across versions but within the middle part the 16 items were arranged into four different random orders.

In the first part (Task One) the participant reported his or her age and sex and indicated whether they recognised any people in the photographic array. The final division (Task Three) comprised of a single 7-point Likert-type scale rating each depicted person's likelihood of being a criminal (CRI; 1- "very unlikely to be a criminal" to 7- "very likely to be a criminal"). Task Two included 16 items, each of which required the participant to rate all 18 photographs. Eight items called for judgments relating to physical qualities of the photographs and their subjects, which had all been manipulated to reduce differences between the SK and non-criminal photographs. These were the depicted person's age (AGE; in full years), grooming (GRO; 1- "very poorly groomed" to 7- "very well groomed"), orientation (ORI; 1- "extreme right" to 7- "extreme left"), babyfaceness (BAB, 1- "very baby-like face" to 7- "very mature face"), expression (EXP; 1- "very negative" to 7- "very positive")

and physical attractiveness (PA; 1- "very physically unattractive" to 7- "very physically attractive"), plus also picture quality (QUA; 1- "very poor" to 7- "very good") and lighting (LIG; 1- "very dark" to 7- "very light").

The remaining eight items required physiognomic judgments pertaining to the personality disorder of psychopathy. These items were based on factor one of the PCL-R (Hare, 1991). The questionnaire items were written to convey the essence of the descriptive prototypical paragraphs against which clinicians using the PCL-R measure the exhibition of psychopathic traits, while condensing them to a level of detail suitable for casual physiognomic estimation by non-clinicians. The items based on the PCL-R were: self-opinion (SEL; 1- "has a very inflated opinion of himself" to 7- "is very modest"), guiltiness (GUI; 1- "does not feel guilt" to 7- "feels guilt very easily"), truthfulness (TRU; 1- "makes a habit of lying" to 7- "makes a habit of telling the truth), manipulative tendencies (MAN; 1- "very manipulative" to 7- "not at all manipulative"), emotional depth (EMO; 1- "emotionally shallow" to "deeply emotional"), attitude of care (CAR; 1- "very uncaring" to 7- "very caring"), sincerity of charm (CHA; 1- "very insincere and superficial" to 7- "very genuine"), and responsibility (RES; 1- "hardly ever" to 7- "nearly always").

Procedure

A main study questionnaire and a main study photographic array were distributed to each participant. The questionnaires were filled out in the presence of the experimenter without communication between participants. Following completion, the answers were checked for missing data and any ambiguous numerals were queried. The participants were also asked to confirm whether they had correctly understood the direction of the Likert-type scales and those who had confused the anchors amended their answers. Instructions were given not to talk to anybody regarding the study until data collection was completed, in an attempt to minimise knowledge of the study's purpose in subsequent groups of participants. The participants were then thanked for their participation, paid and debriefed.

This research was conducted after review by the University of Canterbury Human Ethics Committee.

Results

Overview

Group comparisons between judgments of the sexual serial killer and non-criminal photographs were conducted using analyses of variance. Particular emphasis was given to perceptions of criminal propensity and psychopathic tendency. Analyses of covariance were performed in order to statistically control the influence on perceived criminality and psychopathy that was exerted by those other variables that differed between the two photograph types. Pearson product-moment correlation coefficients and multiple regression analyses were then used to explore the relationship of the variables describing physical qualities of the photographs and their subjects (the physical variables; AGE, GRO, ORI, BAB, EXP, QUA, LIG, and PA) to ratings of criminal propensity and psychopathy over all the photographs. Following this, more group comparisons were made relating to the physical variables through analyses of variance and discriminant function analyses. These comparisons were between those photographs that, according to their type (SK or non-criminal), were either correctly perceived in relationship to psychopathy and criminality or were confused with the other type of photograph. Finally, a discriminant function analyses was performed to distinguish between the SK and non-criminal photographs on the basis of the physical variables.

Group-based Ratings for the Sexual Serial Killer and Non-criminal Photograph Types

The mean ratings and standard deviations for all the measured variables, averaged across participants separately for SK and non-criminal photographs, are shown below in Table 1. Scores for babyfaceness were reverse coded so that a higher score represented a greater level of perceived babyfaceness. Means and standard deviations for all variables for each photograph are reported in Appendix 9.

Table 1. *Mean Ratings for PCL-R derived variables, Psychopathic Traits, Criminality and Physical Variables as a Function of Photograph Type (N=66)*

Variable	SK photographs		Non-criminal photographs	
	Mean	SD	Mean	SD
	PCL-R derived variables			
Self-opinion (SEL)	3.77	0.63	3.87	0.70
Guiltiness (GUI)	3.58_a	0.53	3.96_b	0.63
Truthfulness (TRU)	3.62_a	0.74	4.10_b	0.63
Manipulative tendencies (MAN)	3.72_a	0.55	3.92_b	0.63
Emotional depth (EMO)	3.73_a	0.66	4.13_b	0.66
Attitude of care (CAR)	3.64_a	0.53	4.04_b	0.59
Sincerity of charm (CHA)	3.89_a	0.69	4.21_b	0.63
Responsibility (RES)	3.79_a	0.65	4.27_b	0.63
Main dependent variables				
Psychopathic traits (PSY)	34.26_a	3.61	31.50_b	3.87
Criminality (CRI)	4.50_a	0.62	3.72_b	0.70
Physical variables				
Age (AGE)	36.37_a	3.83	39.45_b	3.70
Grooming (GRO)	4.21_a	0.79	5.03_b	0.57
Orientation (ORI)	4.19	0.42	4.09	0.46
Babyfaceness (BAB)	3.44	0.51	3.52	0.61
Expression (EXP)	3.58	0.51	3.59	0.47
Picture quality (QUA)	4.04_a	0.68	4.26_b	0.64
Lighting (LIG)	3.81	0.52	3.73	0.54
Physical attractiveness (PA)	3.20_a	0.75	3.73_b	0.75

Those means across each row which are labelled by a different subscript are significantly different at $p<.05$ or better according to one-way, within-subject analyses of variance

Perceiving Psychopathic Traits and Criminality

Psychopathic Traits

Reliability coefficients were calculated for the eight scales derived from factor one of the PCL-R, separately for the sexual serial killer and the non-criminal photographs. This yielded a Cronbach's alpha of 0.87 for the SK photographs and 0.90 for the non-criminal photographs. Considering these high reliabilities, the eight scales (SEL, GUI, TRU, MAN, EMO, CAR, CHA, and RES) were aggregated for each photograph type into the single variable of psychopathic traits (PSY), which was reverse coded to facilitate ease of understanding. PSY thus had a possible range between 8 and 56, with a higher score indicating the perception of stronger psychopathic traits than a lower score. This variable was used in all further analyses.

A 2 (sex of participant: female/male) x 4 (photograph order: arrays A/B/C/D) x 4 (question order: questionnaire types I/II/III/IV) x 2 (photograph type: SK photographs/non-criminal photographs) analysis of variance (ANOVA) with repeated measures on the fourth factor was performed for the dependent variable PSY. As predicted, there was a significant main effect for photograph type, $F(1, 34)=39.71$, $p<.0001$, with sexual serial killers being perceived as possessing psychopathic traits to a greater extent than non-criminals ($Ms=34.26$ vs. 31.50). This large effect (Cohen's $f=.78$) was qualified by a significant two-way interaction between photograph order and photograph type, $F(3, 34)=5.40$, $p<.005$, and a significant three-way interaction between sex of participant, photograph order and photograph type, $F(3, 34)=3.04$, $p<.05$. No other effects were significant.

The three-way interaction is displayed in Figure 2. To investigate this interaction, separate 2 (sex of participant: female/male) x 4 (photograph order: arrays A/B/C/D) ANOVAs were conducted for the sexual serial killer and non-criminal photograph types. There were no significant main effects for sex of participant or photograph order, nor was their interaction significant for either the SK photographs or the non-criminal photographs. Because, in the original four-way ANOVA, there were no significant main effects for sex of participant, photograph order or question order and no theoretically suggestive patterns could be discerned in the interactions between sex of the participant, photograph order and photograph type, all subsequent analyses of PSY were collapsed across the between-subject variables.

Criminality

A 2 (sex of participant: female/male) x 4 (photograph order: arrays A/B/C/D) x 4 (question order: questionnaire types I/II/III/IV) x 2 (photograph type: SK photographs/non-criminal photographs) ANOVA with repeated measures for the fourth factor was performed for the dependent variable CRI. There was a significant, large (Cohen's $f=.79$) main effect for photograph type, $F(1, 34)=40.88$, $p<.0001$. In accord with predictions, the mean rating of criminality was higher for the SK photographs than the non-criminal photographs ($Ms=4.50$ vs. 3.72). No other effects approached significance, so means were collapsed across the sex of the participants and order of both photographs and questions in all subsequent analyses of criminality.

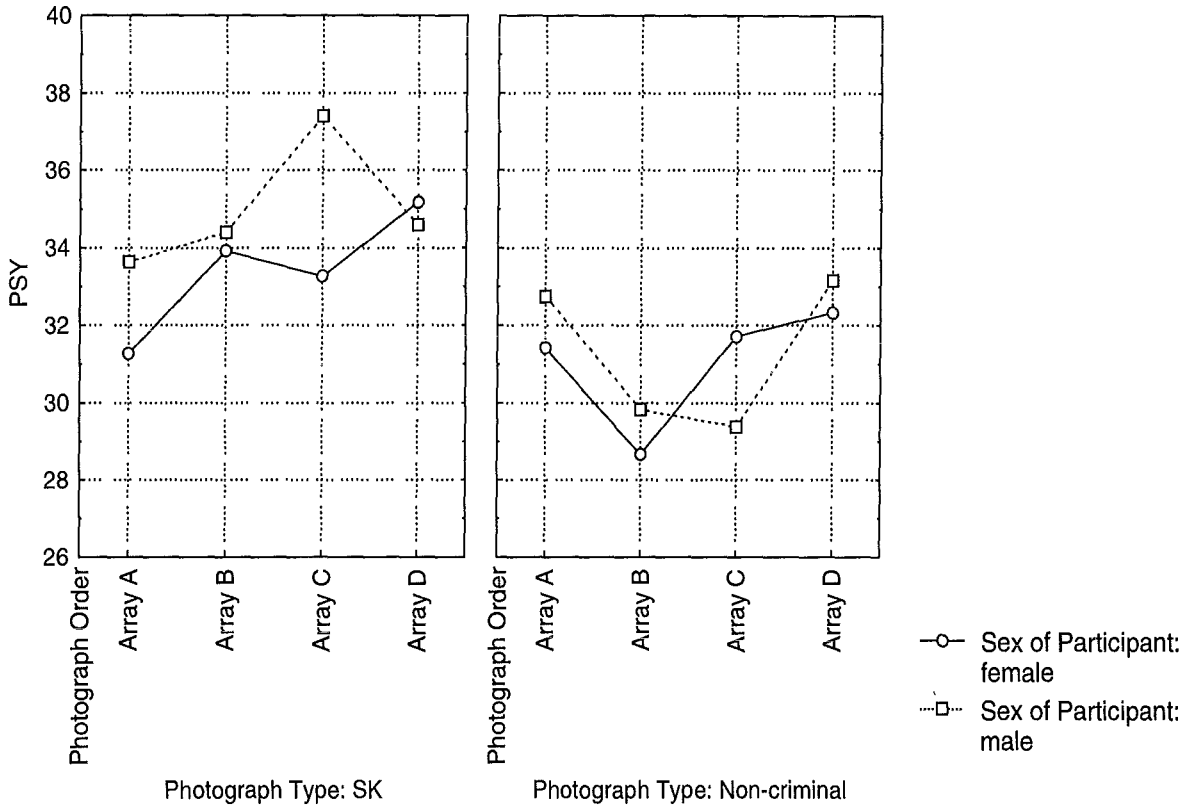


Figure 2. Ratings of psychopathic traits as a function of sex of participant, photograph order and photograph type.

Exploring the Influence of the Physical Variables on PSY and CRI

Further analyses were performed to investigate whether the main effects found for photograph type in ratings of perceptions of the psychopathic traits and criminality could be attributed to differences between the SK and non-criminal photographs on any of the eight physical variables (AGE, GRO, ORI, BAB, EXP, QUA, LIG, and PA).

Controlling Differences in the Physical Variables across Photograph Type

One-way, within-subject ANOVAs were carried out comparing mean ratings of the SK and non-criminal photographs for each physical variable. Despite having

attempted to minimise differences between SK and non-criminal photographs on all of these dimensions, significant differences were found between mean ratings for the estimated age of the targets, $F(1, 65)=97.56$, $p<.0001$, their level of grooming, $F(1, 65)=113.52$, $p<.0001$, picture quality, $F(1, 65)=11.32$, $p<.005$, and their physical attractiveness, $F(1, 65)=45.11$, $p<.0001$. The photographs of the sexual serial killers were perceived as being of lower picture quality (*M*s=4.04 vs. 4.26) and as representing people who were, on average, three years younger (*M*s=36.37 vs. 39.45), less well-groomed (*M*s=4.21 vs. 5.03) and less physically attractive (*M*s=3.20 vs. 3.73) than those of the non-criminal photographs.

These findings raised the possibility that the main effects for photograph type found in perceptions of psychopathic traits and criminality might have been primarily due to the differences between the SK and non-criminal photographs in ratings of age, level of grooming, picture quality and physical attractiveness. To test this hypothesis, a repeated measures analysis of covariance (ANCOVA), with age, grooming, picture quality and physical attractiveness as changing (repeated measures) covariates, was run comparing ratings of the SK and non-criminal photographs for both perceived criminality and perceived psychopathy. The difference between the photograph types was no longer significant for either perceptions of the psychopathic traits, $F(1, 61)=1.98$, $p=.16$, or perceived criminality, $F(1, 61)=1.03$, $p=.31$.

Further ANCOVAs were conducted to gain some indication of which of the variables were primarily responsible for influencing the main effects of photograph type on CRI and PSY. Age, grooming, picture quality and physical attractiveness were used individually as a single covariate of either perceived criminality or the psychopathic traits. For ratings of criminality, only grooming reduced the effect of photograph type below statistical significance by itself. For ratings of perceived psychopathic traits, the effect of photograph type failed to reach statistical significance when any one of age, grooming and physical attractiveness was used as the covariate, but remained significant when only picture quality was controlled.

Relationships between PSY, CRI and the Physical Variables

The results of the ANCOVAs indicated that differences in the physical variables between the sexual serial killer and non-criminal photographs could not be ruled out as being responsible for the main effect of photograph type. Because of this, analyses were conducted to determine to what extent ratings on the physical variables could explain ratings of PSY and CRI.

In order to explore how ratings on the physical variables were related to judgments of psychopathic traits and criminality over the entire range of those photographs, Pearson product-moment correlation coefficients were calculated between the physical variables and both PSY and CRI for mean ratings of each photograph. These are reported in Table 2. Aggregated perceptions of the psychopathic traits were positively correlated with ratings of babyfaceness, $r(18)=.53$, $p<.05$, and negatively correlated with age, $r(18)=-.70$, $p<.001$, and expression, $r(18)=-.59$, $p<.05$. Perceived criminality was negatively correlated with age, $r(18)=-.59$, $p<.01$, grooming, $r(18)=-.54$, $p<.05$, expression, $r(18)=-.72$, $p<.001$, picture quality, $r(18)=-.50$, $p<.05$, and physical attractiveness, $r(18)=-.70$, $p<.001$. A large positive correlation between the main dependent variables, PSY and CRI, was also found between mean ratings for each photograph, $r(18)=.89$, $p<.0001$.

Table 2. *Correlations between Mean Ratings for each Photograph on the Physical Variables and Mean Ratings for each Photograph on PSY and CRI (N=18)*

Variable	PSY	CRI
Age (AGE)	-0.70****	-0.59***
Grooming (GRO)	-0.39	-0.54**
Orientation (ORI)	-0.06	-0.11
Babyfaceness (BAB)	0.53**	0.42*
Expression (EXP)	-0.59**	-0.72****
Picture quality (QUA)	-0.38	-0.50**
Lighting (LIG)	-0.18	-0.34
Physical attractiveness (PA)	-0.45*	-0.70****

* $p<.10$ *** $p<.01$
** $p<.05$ **** $p<.001$

Predicting PSY and CRI using the Physical Variables

To determine the extent to which ratings on the physical variables were predictive of perceptions of the psychopathic traits and criminality, separate multiple regression analyses were conducted on the photographs' mean scores for PSY and CRI. Those physical variables that evidenced at least a marginally significant ($p<.10$) correlation with the dependent variable were entered as independent variables. These were age, babyfaceness, expression and physical attractiveness for PSY and age, grooming, babyfaceness, expression, picture quality and physical attractiveness for CRI.

When the mean ratings of each photograph for PSY were entered as the dependent variable and the photographs' mean ratings for AGE, BAB, EXP and PA, as the independent variables, the regression analysis was significant, $F(4, 13)=7.97$, $p<.005$, and accounted for 62.11% of the variance in the mean PSY scores. Mean perceived age was independently predictive of the mean ratings for perceived psychopathic traits ($\beta=-.99$, $t=-2.68$, $p<.05$). A younger perceived age predicted a higher level of perceived psychopathic traits. The regression analysis was also significant when mean ratings of CRI for the photographs were entered as the dependent variable and the photographs' mean ratings for AGE, GRO, BAB, EXP, QUA and PA were the independent variables, $F(6, 11)=12.82$, $p<.0005$, and accounted for 80.66% of the variance in ratings of perceived criminality. Only mean picture quality was independently predictive of the mean CRI scores for the photographs ($\beta=-.31$, $t=-2.45$, $p<.05$), with the general perception of a photograph as being of poorer picture quality predicting the further shared perception that the person depicted in the photograph was likely to be a serious criminal.

As shown in Table 3, some of the predictor variables were also highly correlated with one another. Age was negatively correlated with babyfaceness, $r(18)=-.91$, $p<.001$, and expression was positively correlated with physical attractiveness $r(18)=.74$, $p<.001$. This led to some ambiguity regarding whether perceived age was independently predictive of perceptions of the psychopathic traits.

Table 3. *Correlations between Mean Ratings for Each Photograph on those Physical Variables used as Predictor Variables in the Multiple Regression Analyses (N=18)*

Variable	AGE	GRO	BAB	EXP	QUA	PA
Age (AGE)	----	0.36	-0.91****	0.25	0.08	0.12
Grooming (GRO)		----	-0.10	0.16	0.37	0.46
Babyfaceness (BAB)			----	-0.17	0.10	0.02
Expression (EXP)				----	0.27	0.74****
Picture quality (QUA)					----	0.21
Physical attractiveness (PA)						----

* $p<.10$ *** $p<.01$
** $p<.05$ **** $p<.001$

Investigating Confusion between the Sexual Serial Killer and Non-criminal Photographs

In an effort to further understand the role that the physical variables might have had in the main effects for photograph type, means for these variables were compared between those photographs that were, on average, correctly classified and those that were mistakenly perceived. The criterion defining confusion in perception was set at the neutral mid-point of the CRI and PSY scales. For instance, it was reasoned that the psychopathic anchors on the PCL-R derived variables (e.g. very manipulative, very shallow) should have been more descriptive of the SKs, and the non-psychopathic anchors (e.g. not at all manipulative, deeply emotional) more descriptive of the non-criminals. Hence, if perceived correctly, SKs should have been rated at over 32 for PSY (the mid-point of the PSY scale range between 8 and 56), this indicating a tendency toward high psychopathic traits. The SK photographs TWO (M=30.86, SD=8.33) and SEVEN (M=28.95, SD=9.00) were said to be mistakenly perceived in relation to psychopathic traits because they had mean PSY ratings of less than 32 while the non-criminals #25 (M=32.32, SD=6.94), #140 (M=43.98, SD=6.66) and #172 (M=34.52, SD=9.10) were misperceived because they were rated over 32. Similarly, for perceived criminality, photographs SIX (M=3.73, SD=1.47), SEVEN (M=2.85, SD=1.56), #140 (M=5.91, SD=1.45), and #176 (M=4.21, SD=1.51) were defined as confused with the opposite type of photographs, based upon their mean CRI ratings and a scale mid-point of four.

Group-based comparisons between photographs categorised according to photograph type and the accuracy of perceptions of the psychopathic traits or criminality. Within-subject ratings relating to the physical variables were averaged separately over the correctly perceived SK photographs, the confused SK photographs, the correctly perceived non-criminal photographs, and the confused non-criminal photographs (categorised separately in relation to mean PSY and CRI scores). Table 4 gives the mean values for the different photograph groups, which were subjected to one-way, within-subject ANOVAs with four levels (photograph type: correctly perceived SK/confused SK/correctly perceived non-criminal/confused non-criminal), followed by separate comparisons between the levels through *post hoc* analyses (Tukey HSD tests).

Table 4. *Mean Scores for the Physical Variables Averaged across Correctly Perceived and Confused Photographs as Defined using either PSY or CRI scores (N=66)*

Variable	SK photographs		Non-criminal photographs	
	Correct	Confused	Correct	Confused
	Categorisations based on PSY scores			
Age (AGE)	35.35_a	39.95_b	43.04_c	32.26_d
Grooming (GRO)	3.89_a	5.31_b	5.19_b	4.70_c
Orientation (ORI)	4.21	4.12	4.01	4.24
Babyfaceness (BAB)	3.46_a	3.38_a	3.17_a	4.20_b
Expression (EXP)	3.11_a	5.23_b	3.70_c	3.36_d
Picture quality (QUA)	4.00_a	4.16_a	4.57_b	3.64_d
Lighting (LIG)	3.64_a	4.42_b	3.78_a	3.63_a
Physical attractiveness (PA)	2.93_a	4.14_b	3.69_c	3.82_c
	Categorisations based on CRI scores			
Age (AGE)	35.07_a	40.92_b	42.94_c	27.25_d
Grooming (GRO)	4.04_a	4.81_b	5.14_c	4.63_b
Orientation (ORI)	4.10_a	4.50_b	4.16_{abc}	3.81_{ac}
Babyfaceness (BAB)	3.66_a	2.67_b	3.08_c	5.04_d
Expression (EXP)	3.13_a	5.15_b	3.88_c	2.58_d
Picture quality (QUA)	4.00_a	4.17_a	4.47_b	3.52_c
Lighting (LIG)	3.71_a	4.18_b	3.79_{ac}	3.50_{ad}
Physical attractiveness (PA)	3.00_a	3.88_b	3.85_b	3.31_a

Note: all one-way ANOVAs are significant at $p<.0005$ or better except for Orientation which is not significant at $p<.05$
Those means across each row which are labelled by a different subscript are significantly different at $p<.05$ according to Tukey's HSD test.

Comparisons were made between: (1) photographs of the same type that were rated differently on PSY or CRI (between the first and second columns and between the

third and fourth columns); and, (2) photographs of different types that were rated similarly on PSY or CRI (between the first and fourth columns and between the second and third columns).

(1) A clear pattern is shown in Table 4 for comparisons between the correct and confused photographs of the same type. The correct and confused photographs within each photograph type (i.e. SK or non-criminal) showed significant differences in ratings for most variables. For the PSY-based groupings of SK photographs, the confused photographs were rated as significantly older (*M*s=39.95 vs. 35.35), better groomed (*M*s=5.31 vs. 3.89), more positive in facial expression (*M*s=5.23 vs. 3.11), lighter (*M*s=4.42 vs. 3.64) and more physically attractive (*M*s=4.14 vs. 2.93) than the correctly perceived photographs. Only orientation, babyfaceness, and picture quality failed to show a significant difference between the correctly perceived and confused SK photographs. For the non-criminal photographs that were grouped on the basis of mean PSY ratings, the confused photographs were rated as significantly younger (*M*s=32.26 vs. 43.04), less well groomed (*M*s=4.70 vs. 5.19), more babyfaced (*M*s=4.20 vs. 3.17), more negative in facial expression (*M*s=3.36 vs. 3.70) and of poorer picture quality (*M*s=3.64 vs. 4.57) than the correctly perceived non-criminal photographs. There were no significant differences between the correctly perceived and confused non-criminal photographs in orientation, lighting and physical attractiveness.

Turning to the CRI-based groupings, the confused SK photographs were perceived to be significantly older (*M*=40.92 vs. *M*=35.07), better groomed (*M*s=4.81 vs. 4.04), more oriented to the left (*M*s=4.50 vs. 4.10), less babyfaced (*M*s=2.67 vs. 3.66), more positive in facial expression (*M*s=5.15 vs. 3.13), lighter (*M*s=4.17 vs. 4.00) and more physically attractive (*M*s=3.88 vs. 3.00) than the correctly perceived SK photographs. There was no significant difference between the correct and confused SK photographs in picture quality only. The confused non-criminal photographs were rated as significantly younger (*M*s=27.25 vs. 42.94), less well groomed (*M*s=4.63 vs. 5.14), more babyfaced (*M*s=2.96 vs. 4.92), more negative in expression (*M*s=2.58 vs. 3.88), poorer in picture quality (*M*s=3.52 vs. 4.47), darker (*M*s=3.50 vs. 3.79), and less physically attractive (*M*s=3.31 vs. 3.85) than the correctly perceived non-criminal

photographs, with only orientation showing no significant difference between the groups.

Of particular interest is that these differences between mean scores for the correct and confused photographs followed the same directions as the correlation coefficients reported in Table 2. For example, perusal of the directions of the relationships in Table 2 (ignoring statistical significance for the time being) shows that, across all 18 photographs, higher perceived psychopathic traits were associated with youth, poor grooming, orientation to the right, babyfaceness, negative expression, poor picture quality, dark lighting and low physical attractiveness. Similarly, misperception of a non-criminal photograph as representing someone with high psychopathic traits was associated with significantly lower age, poorer grooming, more babyfaceness, a more negative expression and lower picture quality than the correctly perceived non-criminal photographs. The same directions of association were also found for the SK photographs, where correctly perceived SK photographs were significantly younger, less well groomed, more negative in expression, darker and less physically attractive than those SK photographs that were confused with the non-criminals. In every case, significant differences between the correctly perceived and misperceived photographs of the same type followed the same direction as predicted by the correlation of mean ratings on the physical variables with perceived psychopathic traits over all 18 photographs. Thus, those photographs that were misperceived in terms of the magnitude of psychopathic traits tended to be more similar, in the physical variables, to the photographs with which they were confused than were the rest of photographs of their same type that were correctly perceived. For example, youth was associated with high mean perceived psychopathic traits over all the photographs, and, consistent with this overall relationship, those non-criminals that were misperceived as high in psychopathic traits were rated as younger than the other non-criminals, while the SKs seen as low in psychopathic traits were rated as older than the correctly judged SKs.

Similarly, for the CRI-based photographic groups, the differences in mean group scores again coincided with the expectations raised by Table 2. Just as for the PSY-based photographic groups, there was not a single variable where the confused and correctly perceived groups of the same photograph type significantly differed in a

direction contrary to the earlier correlations between mean ratings on the physical variables and perceived criminality over all the photographs.

(2) Comparisons between those photographs of different type that were rated similarly on PSY or CRI, that is between the correctly perceived SK photographs and the confused non-criminal photographs and between the correctly perceived non-criminal photographs and the confused SK photographs, did not reveal the same pattern. Again, the different categories tended to show significant differences in ratings for most variables. However, unlike the comparisons between confused and correctly perceived photographs of the same type, comparisons between the photographs that were rated similarly (based on PSY or CRI) but were of different type were not entirely consistent with the relative directions of correlation shown in Table 2. The directions of correlation for relationships between the physical variables and both PSY and CRI over all the photographs (see Table 2) were all the same (negative) except for babyfaceness (positive). Consistency with the overall correlations would have been shown if, for those differences that were significant, one of the groups were higher than the other on all the physical variables apart from babyfaceness, where it would be lower. This did not occur in any of the four comparisons. To give a concrete example, the confused SK photographs were rated higher than the correctly perceived non-criminal photographs (based on PSY) for expression, lighting and physical attractiveness but were, inconsistently, lower on age and picture quality. Similarly, taking another example, the correctly perceived SK photographs were rated higher than the confused non-criminal photographs (based on CRI) for age, expression and picture quality, as well as being rated lower on babyfaceness. However, they were also rated lower on grooming, when, to be entirely consistent with the relative directions of association in Table 2, they should have been rated higher.

For the PSY-based groupings of correctly perceived SK photographs and confused non-criminal photographs, the SK photographs were rated as significantly older (*M*s=35.35 vs. 32.26), more poorly groomed (*M*s=3.89 vs. 4.70), less babyfaced (*M*s=3.46 vs. 4.20), more negative in facial expression (*M*s=3.11 vs. 3.36), of better picture quality (*M*s=4.00 vs. 3.64) and less physically attractive (*M*s=2.93 vs. 3.82) than the non-criminal photographs. Only orientation and lighting failed to show a

significant difference between the correctly perceived SK photographs and confused non-criminal photographs. For the correctly perceived non-criminal photographs and confused SK photographs that were grouped on the basis of mean PSY ratings, the SK photographs were rated as significantly younger (*M*s=39.95 vs. 43.04), more positive in facial expression (*M*s=5.23 vs. 3.70) of poorer picture quality (*M*s=4.16 vs. 4.57), lighter (*M*s=4.42 vs. 3.78), and more physically attractive (*M*s=4.14 vs. 3.69), than the non-criminal photographs. There were no significant differences between the correctly perceived non-criminal photographs and confused SK photographs in grooming, orientation, and babyfaceness.

For the CRI-based groupings of correctly perceived SK photographs and confused non-criminal photographs, the SK photographs were rated as significantly older (*M*s=35.07 vs. 27.25), more poorly groomed (*M*s=4.04 vs. 4.63), less babyfaced (*M*s=3.66 vs. 5.04), more positive in facial expression (*M*s=3.13 vs. 2.58), and of better picture quality (*M*s=4.00 vs. 3.52) than the non-criminal photographs. Orientation, lighting and physical attractiveness failed to show a significant difference between the correctly perceived SK photographs and confused non-criminal photographs. For the correctly perceived non-criminal photographs and confused SK photographs that were grouped on the basis of mean CRI ratings, the SK photographs were rated as significantly younger (*M*s=40.92 vs. 42.94), more poorly groomed (*M*s=4.81 vs. 5.14), less babyfaced (*M*s=2.67 vs. 3.08), more positive in facial expression (*M*s=5.15 vs. 3.88) of poorer picture quality (*M*s=4.17 vs. 4.47), and lighter (*M*s=4.18 vs. 3.79) than the non-criminal photographs. There were no significant differences between the correctly perceived non-criminal photographs and confused SK photographs in orientation and physical attractiveness.

Using the physical variables to discriminate between photographs categorised according to photograph type and the accuracy of perceptions of the psychopathic traits or criminality. Ten forward stepwise discriminant function analyses were performed, in an attempt to find out how mean ratings for each photograph on the physical variables discriminated between: (1) the confused and correctly perceived SK photographs; (2) the confused and correctly perceived non-criminal photographs; (3) the confused SK photographs and the correctly perceived non-criminal photographs; (4) the confused non-criminal photographs and the correctly perceived

SK photographs; (all defined in reference to both PSY and CRI) and, (5) the photographs perceived as more psychopathic in their taits or non-psychopathic and then the photographs perceived as criminal or non-criminal. The criterion required for a variable to enter into a model was set at $F \geq 1.00$ while the criterion for removal was set at $F=0.00$.

(1) The discriminant function analysis using the physical variables to predict membership in the confused or correctly perceived SK photograph groups (defined according to PSY) was significant, Wilks' $\lambda=.21$, $F(2, 6)=11.25$, $p<.01$, and provided a correct classification of 100% of the photographs from a two variable model including physical attractiveness (partial Wilks' $\lambda=.34$, $F(1, 6)=11.47$, $p<.05$, standardised discrimination function coefficient=-.95) and lighting (partial Wilks' $\lambda=.68$, $F(1, 6)=2.85$, *ns*, standardised discrimination function coefficient=-.66). High physical attractiveness predicted that a SK photograph was misperceived as exhibiting low levels of psychopathic traits. The discriminant function analysis predicting confused and correctly perceived SK photographs (defined according to CRI) was also significant, Wilks' $\lambda=.37$, $F(2, 6)=5.18$, $p<.05$, providing a correct classification of 88.89% of the photographs from a two variable model including expression (partial Wilks' $\lambda=.51$, $F(1, 6)=5.84$, *ns*, standardised discrimination function coefficient=-.92) and orientation (partial Wilks' $\lambda=.69$, $F(1, 6)=2.70$, *ns*, standardised discrimination function coefficient=-.73). Photograph SIX was wrongly classified as having been correctly perceived as a SK when, in fact, it was confused with the non-criminals for ratings of CRI.

(2) The discriminant function analysis predicting membership in either the confused or correctly perceived non-criminal photograph groups (defined according to PSY) was significant, Wilks' $\lambda=.01$, $F(6, 2)=32.66$, $p<.05$. It provided a correct classification of 100% of the photographs from a six variable model including age (partial Wilks' $\lambda=.02$, $F(1, 2)=82.17$, $p<.01$, standardised discrimination function coefficient=47.64), babyfaceness (partial Wilks' $\lambda=.03$, $F(1, 2)=61.24$, $p<.05$, standardised discrimination function coefficient=40.81), lighting (partial Wilks' $\lambda=.13$, $F(1, 2)=13.37$, *ns*, standardised discrimination function coefficient=5.64), expression (partial Wilks' $\lambda=.04$, $F(1, 2)=44.76$, $p<.05$, standardised discrimination function coefficient=-10.66), grooming (partial Wilks' $\lambda=.06$, $F(1, 2)=29.03$, $p<.05$,

standardised discrimination function coefficient=-9.12) and picture quality (partial Wilks' λ=.27, $F(1, 2)=5.39$, *ns*, standardised discrimination function coefficient=4.69). Lower age, lower babyfaceness, a more positive expression and better grooming predicted misperception of a non-criminal photograph as exhibiting psychopathic tendencies. The discriminant function analysis predicting confused and correctly perceived non-criminal photographs (defined according to CRI) was also significant, Wilks' λ=.01, $F(5, 3)=43.06$, $p<.01$, although BAB, due to its high correlation with AGE, was removed from the analysis in order to avoid the danger of matrix ill-conditioning. Correct classification of 100% of the photographs resulted from a five variable model including age (partial Wilks' λ=.15, $F(1, 3)=17.39$, $p<.05$, standardised discrimination function coefficient=-2.53), physical attractiveness (partial Wilks' λ=.15, $F(1, 3)=17.12$, $p<.05$, standardised discrimination function coefficient=-3.79), orientation (partial Wilks' λ=.06, $F(1, 3)=44.81$, $p<.01$, standardised discrimination function coefficient=-3.71), grooming (partial Wilks' λ=.16, $F(1, 3)=16.28$, $p<.05$, standardised discrimination function coefficient=-3.20) and expression (partial Wilks' λ=.63, $F(1, 3)=1.78$, *ns*, standardised discrimination function coefficient=-1.17). Misperception of a non-criminal photograph as representing someone who was a criminal was predicted by ratings of lower age, lower physical attractiveness, orientation to the right, and poorer grooming.

(3) When a forward stepwise discrimination function analysis was conducted to distinguish between the confused SK and correctly perceived non-criminal photographs (defined in relation to PSY) on the basis of the physical variables, this analysis was not significant, Wilks' λ=.03, $F(5, 2)=14.42$, *ns*. Once more, BAB was removed from this analysis due to a danger of matrix ill-conditioning caused by its high correlation with AGE. The discriminant function analysis predicting confused SK and correctly perceived non-criminal photographs (defined according to CRI) was significant, Wilks' λ=.05, $F(5, 3)=12.11$, $p<.05$. Correct classification of 100% of the photographs was achieved using a five variable model including orientation (partial Wilks' λ=.77, $F(1, 3)=.92$, *ns*, standardised discrimination function coefficient=-.70), expression (partial Wilks' λ=.09, $F(1, 3)=31.07$, $p<.05$, standardised discrimination function coefficient=-3.36), babyfaceness (partial Wilks' λ=.09, $F(1, 3)=29.07$, $p<.05$, standardised discrimination function coefficient=4.57), lighting (partial Wilks' λ=.14, $F(1, 3)=18.55$, $p<.05$, standardised discrimination function coefficient=-3.07) and

grooming (partial Wilks' λ=.50, $F(1, 3)$=2.99, *ns*, standardised discrimination function coefficient=1.44). Those SKs misperceived as low in criminality could be distinguished from the correctly perceived non-criminals on the basis of possessing more positive expressions and less baby-like faces, and being lighter.

(4) The discriminant function analysis predicting membership in either the confused non-criminal or correctly perceived SK photograph groups (defined according to PSY) was significant, Wilks' λ=.21, $F(3, 6)$=7.37, p<.05, and provided a correct classification of 100% of the photographs from a three variable model including physical attractiveness (partial Wilks' λ=.31, $F(1, 6)$=13.12, p<.05, standardised discrimination function coefficient=1.38), grooming (partial Wilks' λ=.46, $F(1, 6)$=7.18, p<.05, standardised discrimination function coefficient=1.03) and lighting (partial Wilks' λ=.64, $F(1, 6)$=3.37, *ns*, standardised discrimination function coefficient=-.93). Lower physical attractiveness and poorer grooming were predictive of the photographs of SKs correctly rated as having high psychopathic traits as compared to the non-criminal photographs incorrectly rated as highly psychopathic in their traits. The discriminant function analysis predicting confused non-criminal and correctly perceived SK photographs (defined according to CRI) was not significant, Wilks' λ=.33, $F(4, 4)$=2.05, *ns*. Again, BAB was removed from this analysis due to a danger of matrix ill-conditioning caused by its high correlation with AGE.

(5) The discriminant function analysis conducted to distinguish the photographs that were perceived as psychopathic (confused non-criminal and correctly perceived SK photograph groups, defined according to PSY) and those that were perceived as non-psychopathic (confused SK and correctly perceived non-criminal photographs, defined in reference to PSY) was significant, Wilks' λ=.19, $F(4, 13)$=12.96, p<.0005. This analysis gave a correct classification of 100% of the photographs from a four variable model including physical attractiveness (partial Wilks' λ=.45, $F(1, 13)$=16.04, p<.005, standardised discrimination function coefficient=-.94), age (partial Wilks' λ=.58, $F(1, 13)$=9.55, p<.01, standardised discrimination function coefficient=-2.20), babyfaceness (partial Wilks' λ=.77, $F(1, 13)$=3.91, *ns*, standardised discrimination function coefficient=-1.58) and picture quality (partial Wilks' λ=.89, $F(1, 13)$=1.55, *ns*, standardised discrimination function coefficient=-.38). The perception of high psychopathic traits was associated with low physical attractiveness

and youth. The discriminant function analysis predicting perceived criminality (confused non-criminal and correctly perceived SK photographs, defined according to CRI) and those who were not perceived as criminal (confused SK and correctly perceived non-criminal photographs, defined according to CRI) was also significant, Wilks' λ=.10, $F(8, 9)$=10.34, p<.001. Correct classification of 100% of the photographs resulted from an eight variable model including physical attractiveness (partial Wilks' λ=.41, $F(1, 9)$=12.93, p<.01, standardised discrimination function coefficient=1.95), age (partial Wilks' λ=.89, $F(1, 9)$=1.14, *ns*, standardised discrimination function coefficient=-1.05), orientation (partial Wilks' λ=.37, $F(1, 9)$=15.50, p<.005, standardised discrimination function coefficient=1.79), picture quality (partial Wilks' λ=.42, $F(1, 9)$=12.62, p<.01, standardised discrimination function coefficient=2.05, expression, partial Wilks' λ=.88, $F(1, 9)$=1.24, *ns*, standardised discrimination function coefficient=-.63), lighting (partial Wilks' λ=.62, $F(1, 9)$=5.55, p<.05, standardised discrimination function coefficient=-1.34), babyfaceness (partial Wilks' λ=.73, $F(1, 9)$=3.37, *ns*, standardised discrimination function coefficient=-1.71) and grooming (partial Wilks' λ=.74, $F(1, 9)$=3.20, *ns*, standardised discrimination function coefficient=.88). The perception of a photograph as representing a criminal was predicted by lower physical attractiveness, orientation to the right, lower picture quality, and lightness.

Table 5 displays the standardised discriminative function analysis coefficients for those physical variables that significantly discriminated between photograph groups, provided the overall analysis from which they came was also significant. It must be noted that the standardised coefficients are to be interpreted in terms of the particular discriminant functions from which they were derived, and thus their directions are only consistent within each analysis, not between analyses. Broadly speaking, the standardised coefficients reveal a pattern that is supportive of the analysis reported in Table 4 and the correlations between the physical variables and PSY and CRI for mean ratings over all the photographs reported in Table 2. For instance, when interpreted in terms of the relevant discriminant function, lower age predicted those non-criminal photographs that were misperceived as criminal and psychopathic as well as those photographs of both types that were perceived as psychopathic in their traits. High physical attractiveness was associated in general with those photographs perceived as being low in psychopathic traits and low in criminality. More

specifically, high physical attractiveness predicted misperceptions of low psychopathic traits in SK photographs, while misperceptions of high criminality in non-criminal photographs were associated with low physical attractiveness.

Table 5. *Standardised Discrimination Function Coefficients from Discrimination Function Analyses that used Mean Ratings of Photographs (N=18) on the Physical Variables to Predict whether the Photographs were Correctly Perceived or Confused (as Defined using either PSY or CRI scores)*

	Photograph Groups				
Physical Variable	Correct SK Confused SK	Confused Non-criminal Correct Non-criminal	Confused SK Correct Non-criminal	Correct SK Confused Non-criminal	Perceived SK Perceived Non-criminal
	Categorisations based on PSY scores				
AGE	-------	47.64	-------	-------	-2.20
GRO	-------	-9.12	-------	1.03	-------
ORI	-------	-------	-------	-------	-------
BAB	-------	40.81	-------	-------	-------
EXP	-------	-10.66	-------	-------	-------
QUA	-------	-------	-------	-------	-------
LIG	-------	-------	-------	-------	-------
PA	-.95	-------	-------	1.38	-.94
	Categorisations based on CRI scores				
AGE	-------	-2.53	-------	-------	-------
GRO	-------	-3.20	-------	-------	-------
ORI	-------	-3.71	-------	-------	1.79
BAB	-------	-------	4.57	-------	-------
EXP	-------	-------	-3.36	-------	-------
QUA	-------	-------	-------	-------	2.05
LIG	-------	-------	-3.07	-------	-1.34
PA	-------	-3.79	-------	-------	1.95

Note: only those standardised discriminant function coefficients representing variables that both came from a significant discriminant function analysis model and themselves significantly discriminated between the two relevant groups of photographs are displayed. Directions of the coefficients are only consistent relative to variables in the same analysis.

The only exception to this general consistency with the earlier analyses was the comparison between those non-criminal photographs that were rated as high in psychopathic traits and the non-criminal photographs that were correctly perceived according to the psychopathic traits. Contrary to the directions of association shown in Table 2 and the mean group scores detailed in Table 4, the confused photographs could be categorised *post hoc* through using a model that predicted their being less babyfaced, having more positive expressions and being better groomed than the other

non-criminals. However, perhaps more understandably, it was relative youth that discriminated the most between the confused and correctly perceived pictures (as indicated by partial Wilkes' λ). When the discriminant function analysis was repeated without including age as an independent variable, the analysis was not significant, Wilks' λ=.49, $F(2, 6)=3.15$, *ns*. This supported the interpretation that age was the driving force behind the earlier discrimination, and that the anomalous contributions of BAB, EXP and GRO to the function were of less relative importance.

The discriminative function analyses comparing those photographs that were perceived oppositely regarding psychopathic traits (the confused non-criminal and correctly perceived SK photographs versus the confused SK and correctly perceived non-criminal photographs, defined in reference to PSY) and criminality (the same photograph groups, defined in relation to CRI), reported in the final column of Table 5, are generally in agreement with the direction of the correlations in Table 2. Photographs that were rated as high in psychopathic traits could be predicted by low attractiveness and youth, mirroring the directions of the associations between these physical variables and ratings of perceived psychopathic traits across all the photographs. Those photographs rated as criminal were distinguishable from those rated as non-criminal on the basis of lower physical attractiveness, orientation to the right, lower picture quality, and lightness, where only lighting was contrary to the overall trends.

Discriminating between the Sexual Serial Killer and Non-criminal photographs

Having examined the data for relationships between the physical variables and perceptions of criminality and psychopathic traits, a question remained pertaining to what extent the participants' ratings of the photographs on the physical variables could be used to discriminate between the SKs and *prima facie* non-criminal photographs. A forward stepwise discriminant function analysis was conducted with the criterion for entry of a variable into the model set at $F\geq1.00$ and removal at $F=0.00$. The discriminant function analysis using the physical variables as predictors of photograph type (SK or non-criminal) was significant, Wilks' λ=.42, $F(5, 12)=3.30$, $p<.05$, and provided a correct classification of 88.89% of the photographs from a five variable model including grooming (partial Wilks' λ=.96, $F(1, 12)=.50$, *ns*, standardised

discrimination function coefficient=.30), lighting (partial Wilks' λ=.77, $F(1, 12)$=3.68, *ns*, standardised discrimination function coefficient=-.99), physical attractiveness (partial Wilks' λ=.62, $F(1, 12)$=7.49, $p<.05$, standardised discrimination function coefficient=1.74), expression (partial Wilks' λ=.72, $F(1, 12)$=4.61, *ns*, standardised discrimination function coefficient=-1.34) and picture quality (partial Wilks' λ=.87, $F(1, 12)$=1.83, *ns*, standardised discrimination function coefficient=.71). Photographs #115 and #176 were incorrectly classified as SKs. The SK photographs could be distinguished from the non-criminal photographs on the basis of their lower physical attractiveness.

Discussion

The participants demonstrated an ability to accurately rate people portrayed in facial photographs on both the possession of personality traits derived from the first factor of the PCL-R and the likelihood of having engaged in criminal activities. As predicted, those sexual serial killers portrayed in the facial photographs were collectively rated as more likely to be criminal and as possessing psychopathic traits to a greater extent than the group of *prima facie* non-criminals. On a more absolute level, when compared to the neutral midpoints of the PSY and CRI scales, seven of the nine sexual serial killers were perceived as possessing traits that indicated an overall tendency towards psychopathic personality descriptions. Seven out of nine sexual serial killers were also rated as likely to be criminals. These trends were reversed for the *prima facie* non-criminals, with six of nine rated as tending toward non-psychopathic trait descriptions and seven seen as unlikely to be criminals.

Despite pilot testing, the two different types of photographs were rated unequally on perceived age, grooming, physical attractiveness and picture quality. When the covariance of perceived psychopathic traits and criminality with these unequal variables was statistically controlled, the experimental hypothesis was not supported by the data. That is, there was no evidence of the accurate perception of criminal propensity or of personality traits associated with psychopathy that was independent of all differences between the target photograph groups relating to the physical variables. Considering this, it can be said that the research hypothesis was partially supported overall. There was evidence that people could accurately judge criminality and psychopathic traits independent of some, but not all, of the physical variables.

Thus the results allowed the possibility that photograph group differences in ratings of psychopathic traits and criminality might be explicable through other means than physiognomy, such as over-generalisation, although they did not rule out the occurrence of physiognomic perception. In particular, inequalities in perceived age, physical attractiveness, and, most prominently, grooming argued against a conclusion that physiognomy led to the main effect of photograph type on perceptions of the

psychopathic traits. Differences in grooming also challenged isolating physiognomy as the cause of the effect of photograph type on perceived criminality.

Ratings on the physical variables were related to ratings of criminality and the psychopathic traits. This association was shown in analyses across all of the 18 photographs, irrespective of photograph type, and in group-based comparisons between photographs that were assigned to categories on the basis of both real and perceived levels of criminality and psychopathic traits. The group comparisons were most consistent with a relationship linking perceptions of the physical variables to perceptions of psychopathic traits and criminality, rather than to actual levels of psychopathic traits and criminality, as indicated by photograph type.

Over the entire range of 18 photographs, a large proportion of the variance in ratings of perceived criminality and psychopathic traits could be accounted for by the physical variables. Mean criminality scores for the 18 photographs were negatively correlated with perceived age, grooming, positive facial expression, picture quality and physical attractiveness. Perceived criminality could be independently predicted by ratings of picture quality. Mean ratings of the possession of psychopathic traits were negatively correlated with age and positive emotional expression. Babyfaceness was positively related to a perceived tendency toward psychopathic traits. The extent of perceived psychopathic traits was independently predicted by perceived age.

Comparisons were made between those pictures of the same type that were defined *post hoc* as correctly or mistakenly judged. Categorisation was based upon the neutral midpoints of the scales related to criminality and the psychopathic traits and an objective accuracy criterion of photograph type (sexual serial killer or non-criminal). Group differences showed patterns consistent with those associations between the mean ratings of the physical variables, psychopathic traits and criminality that were evident over the entire range of photographs. That is, for means that differed significantly between confused or correctly perceived photographs of the same type, the group that was perceived as higher in criminality or psychopathic traits had a lower mean value for any physical variable that was negatively correlated with criminality or psychopathic traits over all the photographs. Similarly, a significantly higher mean value was found if there was a positive correlation over all the

photographs. In fact, significant differences between the *post hoc* photograph groups of the same type were even consistent with mere trends that had been found over all the photographs. As they involved photographs of the same type, with the same real levels of psychopathic traits and criminal propensity, the comparisons between group means supported a connection between the physical variables and ratings of *perceived* criminality and psychopathic traits.

Discriminative function models were developed to find how the physical variables could be used to accurately categorise photographs of the same type *post hoc*. These models were also largely consistent with the directions of correlation shown in relationships over all the photographs and, again, indicated a connection based on perceived rather than actual levels of criminality and psychopathic traits. For example, the best predictor of mistaken perceptions that a sexual serial killer photograph portrayed someone low in psychopathic traits, as compared to correct perceptions of high psychopathic traits for a sexual serial killer photograph, was high physical attractiveness. Similarly, lower age, lower physical attractiveness, orientation to the right and poorer grooming best discriminated the misperceived non-criminal photographs from those non-criminal photographs that were correctly rated according to criminality.

Group comparisons between photographs of different type provided further support of a broad association between judgments of psychopathic traits, criminality and the physical variables. They indicated that confusion between the photograph types regarding estimations of psychopathic traits or criminal propensity was also related to the participants' ratings of the physical variables. Consistency with the direction of correlations over all of the photographs, found in comparisons between photographs of the same type that were perceived differently regarding the psychopathic traits or criminality, was not found when comparing groups of different type that were perceived similarly regarding the psychopathic traits or criminality. Furthermore, when discriminative function analysis was used to discriminate between these groups on the basis of ratings of the physical variables, it did not prove to be uniformly successful. The physical variables did not significantly discriminate between the confused sexual serial killer and correctly perceived non-criminal photographs (defined in relation to psychopathic traits) and between the confused non-criminal and

correctly perceived sexual serial killer photographs (defined according to criminality). Where the photographs that were confused with each other were significantly distinguishable on the basis of the physical variables, the directions of association were, again, largely consistent with a connection between *perceived* criminality or psychopathic traits and the physical variables. The sexual serial killer photographs that were misperceived as low in criminality could be accurately discriminated from the correctly perceived non-criminal photographs, with which they were confused, through being rated as possessing more positive expressions, less baby-like faces and being lighter. In other words, the sexual serial killer photographs that were confused with the non-criminal photographs were even more typically "non-criminal" in their ratings on the physical variables than were the correctly perceived non-criminal photographs themselves. This pattern can be readily understood in terms of an underlying dimension linking the physical variables to perceived criminality, but makes less sense if actual criminality (i.e. being a sexual serial killer) were associated with negative expressions, babyfaceness and darkness. In the case of the correctly perceived sexual serial killer photographs (defined according to psychopathic traits), these could be discriminated from the misperceived non-criminals on the basis of the former having lower physical attractiveness and poorer grooming. Here the data were equally consistent with low attractiveness and poor grooming being associated with either real or perceived psychopathic traits.

When the photographs that were perceived as high in psychopathic traits, irrespective of type, were compared to those rated as low in psychopathic traits using discriminative function analysis, the former could be discriminated from the latter through their relative youth and physical unattractiveness. Again disregarding photograph type, the photographs perceived as likely to portray criminals could be predicted through being physically unattractive, oriented to the right, light and of poor picture quality, relative to the photographs that were perceived as unlikely to show criminals. Thus, comparisons between photographs grouped by perceptions of psychopathic traits and criminal propensity, rather than by any actual differences, once more showed the same directions of association between the physical variables and psychopathic traits and criminality as were displayed across the entire range of photographs (except for lightness). However, it must be noted that differences in the physical variables could also be used to discriminate between actual differences in

psychopathic traits and criminality. Indeed, ratings on the physical variables could be used to correctly classify 16 of the 18 photographs *post hoc* as either sexual serial killers or non-criminals. Low physical attractiveness independently predicted categorisation as a sexual serial killer, consistent with relationships to ratings of criminality and psychopathic traits over all the photographs.

In summary, the data displayed an association between ratings on the physical variables and judgments regarding criminality and the possession of psychopathic traits. When associations over all the photographs were compared to differences between *post hoc* groups of photographs that were selected in reference to both photograph type and the accuracy of ratings on psychopathic traits or criminality, there was support for the conclusion that the overall trends related the physical variables to *perceived* rather than *real* levels of psychopathic traits and criminality. That is, when photographs of the same type were sorted into two groups based on high or low perceptions of psychopathic traits or criminality, the groups also differed in the physical variables in directions that were consistent with the associations over the entire range of photographs. When photographs of different types that were perceived similarly in relation to psychopathic traits and criminality were compared regarding the physical variables, differences did not follow any consistent direction and significant discriminative functions could not be identified in every case to distinguish between the groups. Thus confusion in perceptions of criminality and psychopathic traits seemed to go hand in hand with photographs of different type that were indistinguishable regarding the physical variables. When similarly perceived photographs of different type could be distinguished on the basis of the physical variables, the differences were still consistent with perceived rather than real character. Finally, when all the photographs that were perceived as high in psychopathic traits or criminal propensity were compared with those low in these perceptions, membership in the groups could be predicted by differences in the physical variables which, again, followed the same directions as the overall associations between higher and lower levels of the physical variables and higher or lower ratings on criminality and psychopathic traits.

Implications of the Results

Accurate Facial Judgments

The results of this study supported accuracy in judgments of criminal propensity and the possession of psychopathic traits from facial photographs. Thus earlier findings of an ability to accurately estimate the particular offences of a person's past from his face (Kozeny, 1962 cited by Bull & Green, 1980; Thornton, 1939) were extended to accuracy in judging general criminal propensity. Likewise, the results of Cherulnik et al. (1981), indicating accurate discriminations between high and low Machiavellian men from facial photographs, were extended to evaluations of those personality traits associated with the first factor of the PCL-R.

In addition, this study was the first in which an attempt was made to control some of the features that could form the basis of accurate perception and, also, the first where the participants rated the stimuli on various broadly described physical characteristics that might be related to accuracy. As, despite efforts at control, some of these variables were found to vary between the photograph types, the results of this study allowed some investigation of the cues that enable accurate judgments of criminal propensity and psychopathic traits from facial photographs. Orientation, lighting, expression and babyfaceness were, as intended for all the variables, successfully equalised across the two photographic groups in this study. Hence these variables were not implicated as possible explanations for the disparate ratings on criminality and psychopathic traits between the sexual serial killer and non-criminal photographs. While it was established that they were not necessary for accurate perception, the present results cannot be taken as proof that these variables are unrelated to accurate judgments of character. Any such relationship was not tested in the present study for these four variables.

There was some indication that significant differences between the photograph types regarding certain physical characteristics of their subjects could have formed the basis of accurate ratings of criminality and psychopathic traits. For perceptions of criminality, covariance with grooming could not be statistically precluded as an independent cause of the different ratings for the photograph types. For ratings of

psychopathic traits, imbalances in age, grooming and physical attractiveness could not be ruled out as independent causes of the main effect of photograph type.

In particular, perceived age independently predicted ratings on psychopathic traits over the 18 photographs and was also implicated as relevant to judgments of psychopathic traits by group comparisons between those photographs that were defined as correctly or incorrectly perceived. These findings underline the possible importance of an influence of age on judgments of psychopathic traits, an influence which might explain their accuracy. However, any importance of age in accurately perceiving psychopathic traits from facial photographs should be read in light of the apparent stability of factor one of the PCL-R over the lifespan (Harpur & Hare, 1994). Because age does not appear to mellow the psychopathic personality, it seems probable that perceived age was an accurate indicator of psychopathic traits in this study only because of a difference in perceived age that was peculiar to those sexual serial killers and non-criminals who were selected. Age would not be expected to accurately distinguish between people in a sample that only included offenders, either psychopathic or non-psychopathic.

Physical attractiveness was not as clearly related to perceptions of psychopathic traits as age, but there was a pattern indicating its relevance. There was a marginally significant negative relationship with ratings on psychopathic traits over all 18 photographs, although this was muddied by a large correlation between physical attractiveness and expression, which itself negatively correlated with the psychopathic traits. In addition, the sexual serial killers who were incorrectly perceived as low in psychopathic traits were significantly more attractive than those killers who were correctly judged. A difference in physical attractiveness could be used to discriminate between these groups *post hoc*, as well as between those targets perceived as high or low in psychopathic traits irrespective of photograph type. Thus the putative role of physical attractiveness in accurate judgments of psychopathic traits, allowed for by the analyses of covariance, was supported by some links between ratings of physical attractiveness and psychopathic traits. This was consistent with both the literature indicating that offenders are less physically attractive than their less antisocial peers (Agnew, 1984; Cavior & Howard, 1973) and the attractiveness halo (e.g. Saladin et al., 1988). Differences in perceived physical attractiveness between sexual serial

killers and more attractive non-criminals might have led to differences in judgments regarding psychopathic traits, however, the direction of any causative relationship between physical attractiveness and judgments on the psychopathic traits was ambiguous.

It was also not possible to rule out that differences in grooming between the sexual serial killers and non-criminals explained the differences in ratings of both psychopathic traits and criminality. That is, a bias linking poor grooming to crime and psychopathic traits, combined with a perception that the sexual serial killers were less well groomed than the non-criminals, might account for why the sexual serial killers were rated as more likely to be criminal and as possessing more psychopathic traits. Furthermore, evidence for a link between grooming and perceptions of both criminality and psychopathic traits was shown in the results. Grooming was negatively related to criminality across the 18 photographs while, for group comparisons between photographs of the same type, the photographs perceived as high in criminality or psychopathy were uniformly less well groomed than those that were perceived as low in criminality or psychopathic traits. The results were consistent with accuracy arising from participants associating poor grooming with bad character. Such an association might have contributed, for example, to earlier findings that low physical attractiveness was related to aspects of criminality, where grooming was either explicitly included in ratings of attractiveness or was not controlled (e.g. Agnew, 1984; Cavior & Howard, 1973; Stewart, 1980, 1985; Zebrowitz, Andreoletti et al., 1998; Zebrowitz & Lee 1999).

In summary, the present results qualify previous research by suggesting that accuracy in judging character from facial photographs may be linked to the perception of differences in certain specific physical characteristics of the photographs and their subjects. In particular, indications were found that perceptions of age, grooming and physical attractiveness were linked to accuracy in judgments regarding psychopathic traits and grooming was linked to accurate judgments of criminality.

Physiognomy

The aim of this study was not only to demonstrate accuracy in judgments from facial photographs but to isolate physiognomy as the mechanism for that accuracy. Global facial descriptions such as age, physical attractiveness, babyfaceness and expression could have relevance to facial judgments of criminal propensity and the possession of psychopathic traits at more than one level. That is, these variables might not only be linked to structural and expressional physiognomy but also to cognitive processes like temporal extension from transient facial expressions or to stereotypes such as the attractiveness halo. These, in turn, might lead to self-fulfilling prophecy effects and a real appearance-trait relationship. Alternatively, they could constitute biasing processes operative at the time of judgment, processes which would have little relevance in real-world relationships between criminal propensity and facial appearance but which might explain apparent findings of accuracy if the stimulus photographs happened to differ in characteristics that were relevant to a bias. For example, the particular sexual serial killers in this study were less attractive than the *prima facie* non-criminals. Even if there is no real appearance-trait relationship between physical attractiveness and criminality, a bias such as the attractiveness halo had the potential to lead to a misguided finding of accuracy in perceiving criminality.

Compounding this inherent ambiguity, the photographs in this study were not randomly selected from the entire population of sexual serial killers and non-criminals and so any systematic differences between the groups were not necessarily representative of the populations of sexual serial killers and non-criminals. It was likely, for instance, that the sexual serial killers were not unduly happy when they had their photographs taken following their arrests, while the portrait anthologies used to obtain non-criminal photographs tended to include the older, more attractive, more successful/happy and better groomed. Thus random selection could not be relied upon to equalise group differences in those possibly confounding physical variables that are exhibited in the same facial features that allow physiognomy. Consequently, the possibility of accurate perception (i.e. group differences in the psychopathic traits and criminality) being due to confounding differences between the photographs interacting with biases operating at the time of judgment, as well as self-fulfilling prophecy effects, was recognised and incorporated into the design. This was done by

equalisation of many physical features of the photographs and their subjects across the photograph types.

However, contrary to the intent behind this study's design, it was impossible to rule out perceptions of certain physical characteristics of the photographs and their subjects as causes of the differences between photograph types in ratings of psychopathic traits and criminality. As mentioned previously, the aim of finding support for physiognomy by controlling the influence of non-physiognomic confounds and variables associated with stereotypes was frustrated, at least partially, by differences that were found between photograph types regarding ratings of some of these variables. Thus, despite findings of accuracy, the research hypothesis was only partially supported. However, the results did provide stronger support for physiognomy than previous studies. They revealed that accuracy in facial judgments could occur independently of orientation, babyfaceness, expression and lighting. Thus, in particular, over-generalisations from babyfaceness or expression did not explain accuracy by themselves, reducing the number of alternative explanations for the accuracy that was found and slightly boosting the feasibility of a physiognomic explanation.

Why was the research hypothesis not completely supported? One possible answer revolves around the global nature of descriptors of facial appearance such as physical attractiveness and emotional expression. It is quite probable that valid cues related to structural and expressional physiognomy are subsumed by impressions of age, babyfaceness, physical attractiveness and expression. For example, eyes that are distinctively close together might, for argument's sake, be linked to both low attractiveness and criminality. Hence, controlling these physical variables, either by experimental manipulation or through statistical means, might have prevented an over-representation of those cues among the criminal and psychopathic that could have been behind earlier findings of accurate facial judgment. In other words, earlier studies could have involved valid physiognomic judgments while the present design controlled the very invariants that such physiognomic judgments are based upon. This possibility was inherent to the experimental design, an inevitable consequence of attempting to support physiognomy through ruling out major alternative explanations

based on stereotyping and other over-generalisations. As ever, failure to rule out the null hypothesis cannot be taken as proof that it is true.

Alternatively, the flip side of the above interpretation is that the results of some earlier studies, despite appearances to the contrary, did not actually support physiognomic perception because the researchers failed to equalise the target photographs across confounding variables. Failure to take adequate steps to control confounds is a criticism that could be levelled at much of the literature involved with accuracy in physiognomic judgments. In most studies the target photographs were randomly, or at least non-systematically, selected. Thus, following the logic of null hypothesis significance testing, any differences over and above those attributable to the manipulations should be accounted for over the long run of experimentation. However, this statistical protection does not extend to those differences that are themselves systematically linked to the dependent variables.

For example, a method of controlling emotional expressions that was used in many previous studies is instructing the target to maintain a neutral face while the stimulus photograph is taken. The study run by Malatesta et al., (1987) has been interpreted to demonstrate that "neutral" expressions can be "mistaken" by judges as expressions of emotion. An alternative view to the interpretation that judges mistakenly impute emotion where there is none might be that participants can judge underlying emotions that are actually being experienced at the time a photograph is taken, despite efforts by the targets to follow instructions to dampen their facial expressions. If this position is correct then many of the studies taken to support accurate physiognomic judgments might be explicable through perceptions of subtle transient emotions that co-vary with the dependent variable and the inference process of temporal extension. For instance, it might be thought that aspects of personality such as power and warmth (Berry, 1990, 1991) would be vulnerable to this possibility.

The challenge of uncontrolled confounds is most obvious for those studies using membership in pre-existing categories as the accuracy criterion. While many permanent differences between the groups, such as physical attractiveness or babyfaceness, could be viewed as proper bases for accurate physiognomic judgments, the greatest danger to valid interpretations would come from transient differences

such as contemporaneous emotional expression, grooming, environmental traces, and characteristics of the photography, like lighting, quality and orientation of the face.

Thornton (1939) used photographs that were taken at incarceration, allowing the possibility that the accurate judgments found regarding the targets' criminal offences were based on, say, tattoos, or hairstyles, or scars or bruises that may have been more prevalent among the violent offenders. It is not even entirely unfeasible that people who committed different crimes might be feeling different emotions at the time of incarceration and that instructions to keep a neutral expression while photographs were taken would have been even less effective in this context than in the laboratory. It is unlikely that such instructions *completely* muted a degree of negative emotional response to the situation. Meanwhile, Cherulnik et al. (1981), who supposedly demonstrated accuracy in judgments of Machiavellianism, used still shots taken from a videotaped interview and did not report any efforts to control factors such as personal grooming, or even the clothing worn by the targets. (It is uncertain from their report exactly how much of the targets' bodies were shown to the participants.) The targets' faces were said to show "no distinct facial expression" (p. 390) but this did not preclude the probability that they were experiencing emotion during the interview, nor the possibility that this emotion was perceivable despite its lack of distinctiveness. Thus, another possible interpretation of the results of the present study, compared to those earlier, would be that the apparent accuracy found in the previous research was merely an artefact of confounding, impermanent differences between stimulus groups, differences that the present design attempted to control.

If, in fact, previous findings of accuracy relating to crime and Machiavellianism were the result of lack of control regarding stimuli differences that were impermanent or irrelevant to physiognomy, then this would have some importance in assessing the evidence supporting the predictions of an ecological approach to facial perception. For a start, the other traits that have been shown to be accurately perceivable, without exception, describe the quality of interpersonal interactions within a social context. This was taken to form the beginning of a pattern regarding the type of traits that are perceivable from the face, one that supported an ecological approach to person perception based on affordances. The accurate judgments of criminality and psychopathic traits found in the present study seem to help to add substance to that

pattern. However, traits relating to interpersonal interactions are also probably particularly vulnerable to the confounding effects of transient physical variables like grooming and emotional expression. This raises the possibility that all previous indications of physiognomic accuracy have actually been misleading. Such a conclusion would be consistent with the studies that found no support for the validity of physiognomy regarding traits that are not social in essence (e.g. Cook, 1939; Gurnee, 1934; Ray, 1958) and would not necessarily contradict an ecological approach to person perception. After all, an ecological approach does suggest that more information would be available in dynamic events such as emotional expression than in the static displays of fixed facial features (McArthur & Baron, 1983; Zebrowitz & Collins, 1997).

If, for argument's sake, physiognomic validity were taken for granted, then there is nothing that obviously distinguishes between perceptions of criminality and psychopathic traits and those traits that have been found to be accurately perceived. The traits, other than Machiavellianism and criminal offences, for which accurate physiognomic perception has been supported can be roughly grouped as relating to either social dominance (Berry, 1990, 1991; Cherulnik et al., 1981; Mazur et al., 1984) or social desirability, that is, characteristics related to antisocial behaviour such as honesty/dishonesty, warmth/coldness and kindness/cruelty (Berry, 1990, 1991; Bond et al., 1994; Zebrowitz et al., 1996). While social dominance might be argued to be more related to reproductive fitness than indications of antisocial tendencies, and therefore perhaps constitutes a more vital affordance in evolutionary terms, there is little to distinguish criminality or the psychopathic traits from the traits listed above as relating to social desirability. Thus, according to an ecological approach, there is no obvious reason to think that physiognomy was involved in those earlier studies finding accuracy in facial perceptions of social desirability if evidence cannot be found to support physiognomy for perceptions of psychopathic traits or general criminal propensity. As a consequence, an interpretation that apparent support for physiognomy in the past, at least in judging social desirability, is simply an illusion caused by the interaction of biases and uncontrolled confounds would appear to be most consistent with the results of the present study.

In conclusion, it is easy to speculate why this study did not produce differences between the sexual serial killer and non-criminal photographs in ratings of criminality and the extent of psychopathic traits that were clearly independent of the physical variables, but only further research can directly assess the possibilities raised by such speculation. The results of this study stress the danger that apparent findings of physiognomic accuracy in some research might be largely attributable to biases operative at the time of the judgments of participants, which interact with theoretically irrelevant differences between the photographs such as levels of grooming. The data also raise the possibility that any legitimate permanent facial cues that may exist indicating the presence of psychopathic traits might also contribute to perceptions of broad facial descriptions such as age and physical attractiveness. These possibilities imply that differences between photographs on these variables should be taken very seriously, both as possible confounding factors and as the agents of true physiognomic discriminations.

Cognitive Explanations for Consensus

As mentioned earlier, a number of previous studies have been interpreted to support the validity of facial stereotypes for criminals (e.g. Bull & Green, 1980; Goldstein et al., 1984; Shoemaker et al., 1973; Yarmey, 1993) but none of these has investigated the actual features that comprise these stereotypes. In fact, only Secord et al. (1954) have made an extensive attempt to describe the particular facial features connected to attributions of personality, but they did not focus on criminality or psychopathic personality traits. Whereas there was incomplete support for accurate physiognomic perception independent of differences between the photographic groups in the physical variables, this study provides considerable support for an association between the physical variables and perceptions of the extent of psychopathic traits and criminality. The present results connect attributions of criminality and psychopathic traits to general descriptions of facial appearance and therefore provide information that contributes to an understanding of what constitutes a stereotypically criminal face at a descriptive level that is somewhat broader than the specific features measured by Secord et al. (1954).

Unfortunately, this study was not primarily designed to explore these associations, but rather to control their influence. This resulted in an inevitable ambiguity regarding the directions of causation, as well as a specific difficulty in ruling out actual criminality and psychopathy as third factors that explain correlations between ratings of psychopathic traits, criminality and the physical variables. Nonetheless, the data were still consistent with processes of over-generalisation because, broadly speaking, they revealed a connection between *perceptions* of criminality and psychopathic traits and ratings of the photographs on the physical variables. Thus this study provides hints regarding the stereotypical appearance of criminals and those with psychopathic traits. In particular, the results suggest that the faces of criminals are believed to be youthful in appearance (relative to a range of ages between 24 and 52), poorly groomed, negative in emotional expression and physically unattractive. People who have high levels of psychopathic traits are pictured to be relatively youthful, babyfaced and negative in their expressions.

Over-generalisation from adaptive perceptions. Zebrowitz (1997) contends that the evolutionarily adaptive function of determining certain characteristics from faces has lead to such a preparedness to utilise facial cues to these attributes that our responses to them have been over-generalised to inappropriate situations. The physical variables of age, physical attractiveness, babyfaceness and emotional expression are all examples of variables that Zebrowitz (1997) connects with the adaptive perception of another's health, maturity and emotional state. She regards the likely consequences of over-generalisation based upon these physical variables as lesser evolutionary evils than a failure to respond to such cues entirely. The results of this study constitute evidence that is consistent with the over-generalisation of adaptive perception and support Zebrowitz' (1997) argument that these form the basis of reading faces.

Physical attractiveness has been hypothesised to indicate youth and health, key markers of evolutionary fitness (Zebrowitz, 1997), and research into the attractiveness halo has provided many examples of over-generalisation to psychological attributes (see Feingold, 1992). For instance, low physical attractiveness has been linked to attributions of criminality (Saladin et al., 1988), psychopathology (Cash, Kehr, Polyson, & Freeman, 1977; Jones, Hansson, & Phillips, 1978) and social deviance (e.g. Dunkle & Francis, 1996; Unger, Hilderbrand, & Madar, 1982). Consistent with

this literature, there was a general pattern in the present data negatively linking physical attractiveness to perceptions of psychopathic traits and criminality. Physical attractiveness was negatively associated with perceived criminal propensity across all the photographs, as well as being significantly different for three of four group comparisons between correctly and incorrectly perceived photographs of the same type, and an important factor in five out of eight significant models discriminating between photographs on the basis of photograph type and the accuracy of perceptions according to psychopathic traits and criminality. This fits a process of over-generalisation from physical attractiveness to psychological traits and social behaviour.

Similarly, age, which is another indicator of fitness and, in reality, closely linked to the frequency of the commission of criminal acts (see Greenberg, 1996) but which has less effect on scores for factor one of the PCL-R (Harpur & Hare, 1994), was negatively correlated with criminality and the psychopathic traits. Age was uniformly lower in those photographs of the same type that were perceived as likely to be criminal or psychopathic in their traits than in those that were rated as unlikely to be criminals or more non-psychopathic in their traits. It was also a significant factor in three of the eight significant models discriminating between photographs grouped on the basis of photograph type and the accuracy of perceptions according to ratings of psychopathic traits and criminality. The data were therefore consistent with over-generalisation from perceived age to criminality and psychopathic traits.

Alternatively, the participants could have been estimating criminality using knowledge that the real frequency of criminal acts peaks in the late teens, or, in terms of the age range of the photographs in the present study, diminishes as age increases. The scales derived from the PCL-R might also be considered as inherently linked to notions of maturity; that is, immodesty, freedom from guilt, lying, manipulation, shallow emotions, lack of care, and irresponsibility might describe the young just as aptly as they denote psychopathy. However, this last possibility does not really argue against over-generalisation from age, it merely highlights the particular relevance of the psychopathic traits to measuring any over-generalisation that may have occurred.

The positive nature of facial expressions was negatively correlated with both perceived criminal propensity and the perceived extent of the psychopathic traits. Like age, mean expression was, without exception, more negative for the groups of photographs that were rated as likely to be criminal or more psychopathic in their traits when compared to the photographs of the same type that were rated as unlikely to be criminals or more non-psychopathic. Although, despite this, expression did not play a large part in *post hoc* discriminations between the categories of photographs based on photograph type and the correctness of perceptions of psychopathic traits and criminality, there is enough evidence to suggest that the participants' ratings of expression were consistent with a process of over-generalisation from adaptive perceptions. In particular, the participants may have been deceived by the positive expressions of the sexual serial killers that they mistakenly rated as unlikely to be criminals; over-generalising from facial cues indicating positive emotions and thus temporarily increased safety in social interactions to false perceptions of permanently positive, safe, non-criminal personalities.

The only physical variable that appeared to directly contradict a process of over-generalisation from evolutionarily adaptive perceptions was babyfaceness. Whereas babyfaceness is hypothesised to encourage protective behaviour toward infants in adults (Berry & McArthur, 1985), which has been over-generalised to attributions of infantile traits in babyfaced adults, including a lack of antisocial tendencies (Zebrowitz, Collins, & Dutta, 1998), the results of this study linked babyfaceness to criminality and higher levels of the psychopathic traits. This does not support a process of over-generalisation from an adaptive perceptual attunement. However, there is good reason to believe that these results may have come about due to a defect in the scale measuring babyfaceness and an insufficient explanation of the babyface construct to the participants. The scale that was intended to measure perceptions of babyfaceness merely included the opposing anchors of "This person has a very baby-like face" and "This person has a very mature face." There was no mention of what babyfaceness or facial maturity were meant to denote and no indication that these descriptions were independent of the age of the targets. This was a mistake, as it seems that the participants probably took a mature face or a baby-like face to be roughly synonymous with an old or young face, respectively. It is noteworthy that babyfaceness was negatively correlated with age to an extremely high degree,

strongly supporting the argument that the participants confounded ratings of babyfaceness with perceptions of age. In retrospect, babyfaceness should have been briefly defined in the preliminary and main study questionnaires, and the distinction from age clarified.

Possible confounds. The physical variables of grooming, orientation, picture quality and lighting, measured differences between the stimuli that were thought might affect ratings of psychopathic traits and criminality but were irrelevant to physiognomy. The data revealed an interesting pattern regarding their influence on perceptions of the psychopathic traits and criminal propensity. There was a suggestion that, when told that the array contained photographs of two different types of people, serious criminals and non-criminals, the participants may have been more vigorous in seeking out any perceivable difference between the photographs in an attempt to sort the criminals from the non-criminals. In contrast, when they were simply rating a group of photographs on personality traits they placed more relative importance on the characteristics of the people portrayed. That is, none of grooming, orientation, picture quality and lighting was significantly associated with ratings of psychopathic traits across all the photographs, but grooming and picture quality were both negatively correlated with perceived criminality. Furthermore, the photographs that were perceived as likely to portray criminals could be distinguished from those thought to represent non-criminals using a model including significant contributions from orientation, picture quality and lighting (plus physical attractiveness). None of the non-physiognomic physical variables provided a significant contribution to the discriminant function analysis distinguishing between the targets that were rated high or low on psychopathic traits.

When considering the variables individually, there is little evidence to indicate that the direction the target was looking affected the participants' perceptions. Orientation was not significantly correlated with either psychopathic traits or criminality across all the photographs. In comparisons between the correctly and incorrectly perceived photographs of the same type, orientation only differed significantly between the sexual serial killer photographs where confusion was defined according to criminality scores. This, on its own, does not constitute a consistent pattern linking orientation to perceptions of psychopathic traits or criminality. Lighting, also, was not significantly

correlated with psychopathic traits or criminality across the entire range of the photographs, but in three of four comparisons between confused and correctly perceived photographs of the same type, perceptions of likely criminality or high psychopathic traits were linked to significantly darker photographs, consistent with the direction of a trend in the correlations. So, there is some suggestion that the lighting of photographs might affect participants' perceptions. Likewise, there are indications that picture quality may have been used to rate psychopathic traits and criminality. Picture quality was negatively correlated to perceptions of criminality across all the photographs, and was significantly lower in the confused, compared with correctly perceived, non-criminal photographs (defined according to both psychopathic traits and criminality). Grooming showed the most consistent associations. Grooming was negatively correlated with criminality across all the photographs and this relationship was reflected in significant differences between the groups of photographs of the same type that were perceived correctly or incorrectly according to ratings of both psychopathic traits and criminality. In all four comparisons, the photographs perceived as more criminal or more psychopathic in their traits were also poorer in their grooming.

Cognitive inference processes. Some of the results could be readily described in terms of the cognitive inference processes that Secord (1958) proposed as bridges between facial features and judgments of personality. For example, the data relating to the positive expressions of the SK photographs that were confused with the non-criminals, raised above in support of an emotion over-generalisation effect, might also be said to indicate the specific inference process of temporal extension. Similarly, the connection between darkness and crime might be viewed as metaphorical, and the link between youth and psychopathy could be described as an example of categorical inference. However, the design of this study does not speak directly to the specifics of the inferences that the participants may have made and the ease with which these processes can be used to explain the results cannot be taken as support for their actual involvement.

Limitations of the Results and Directions for Future Research

Generalisability of the Results

It could be argued that people other than students might have greater abilities of physiognomic perception regarding criminality and psychopathic traits because they are more attuned to the particular stimulus invariants that indicate these affordances (McArthur & Baron, 1983). For example, the perceptual learning histories of those who work in the justice system might provide advantages in accuracy over university students who have less experience with people who are antisocial. However, an ability to accurately perceive dangerousness could be expected to be of such adaptive advantage that, if present at all, it should reside to some degree in most people. While students might not be as accurate as policemen, any accuracy in policemen that is partially attributable to processes of evolution presupposes a basic capacity that should be shared with other humans. Furthermore, specialist experience has not been shown to be of advantage in previous studies of physiognomic perception. Kessen (1957) failed to discover any differences between undergraduates, parents and professional psychologists in their levels of accuracy when rating the personalities of children; nor has the extent of relevant experience held been found to affect the accuracy of judges estimating intelligence from adult faces (Cook, 1939; Ray, 1958). It is also worthy of note that policemen appear to share the same conceptions as the general public regarding the faces that fit particular offences (Bull & Green, 1980), suggesting that they utilise similar cues for facial judgments. When all is considered, there seems to be no reason to suspect that undergraduate students would not represent a basic capacity of physiognomic judgment that might be shared in the general population. What is more, there is not even any evidence to show that other groups of people might differ in the extent to which they have developed that hypothesised capacity.

Assumptions regarding Real Levels of Criminal Propensity and Psychopathic Traits

Weaknesses in the assumptions underlying the experimental manipulations might explain the failure to unequivocally support physiognomic perception. It is possible that the use of sexual serial killers to represent high levels of criminal propensity and

psychopathy was inappropriate. For instance, it might be argued that sexual serial killers are not typical offenders and do not truly demonstrate a general propensity toward crime, instead exhibiting a specific propensity for sadistic murder. However, these people share a willingness to commit extreme crimes that might be expected to incorporate tendencies towards lesser antisocial acts. Furthermore, they do, in fact, share histories of less serious antisocial behaviour persisting from childhood (Geberth & Turco, 1997), and have often been arrested previous to conviction for murder (Warren et al., 1996).

A more serious objection might be that a diagnosis of psychopathic personality disorder was not definitively established for each of the sexual serial killers, nor were criminality and psychopathy ruled out for the *prima facie* non-criminals. Arguably, the degree of real difference between the groups may not have been sufficient to allow a finding of physiognomic accuracy. However, considering the low prevalence of psychopathy in the general population and the wide assumption in the literature that it is almost universal among sexual serial killers, plus the extreme offending evident among the sexual serial killers and the apparent respectability indicated by the occupations of the majority of the non-criminals, any argument that there was a lack of real contrast between the photograph types seems to lack persuasive force.

Future research might address these, fairly minor, concerns by collecting photographs from people who have been convicted of more mundane offences and who have been diagnosed as having psychopathic personality disorder. Of course, this is more easily said than done. As noted by Bull and Rumsey (1988), one of these two researchers once attempted to replicate the study by Thornton (1939), where the particular offences committed by those incarcerated in prison were accurately estimated from their photographs, but was forced to desist from his intentions because he was simply "not given permission to use official photographs of convicted individuals" (p.83).

Inequalities between the Photograph Types in the Physical Variables

As discussed previously, it was intended to control certain physical qualities of the photographs and their subjects that could provide non-physiognomic explanations for any differences in ratings on psychopathic traits and criminality given to each group

(sexual serial killers or non-criminals). The single most important limitation to the results of this study was the ambiguity introduced by failure to control age, grooming, picture quality and physical attractiveness across the photograph types.

While the data from this study cannot be said to prove the variables of grooming, picture quality and lighting affect perceptions from facial photographs, they do raise a warning that seemingly small and irrelevant differences between photographs might be behind any distinctions drawn by participants. It should be stressed that these non-physiognomic variables were subjected to manipulations that reduced the variation between photographs. For instance, pictures of particularly poor quality were enhanced using the "blur" and "sharpen" functions in Adobe Photoshop 4.0 and all the photographs were manipulated to a similar level of darkness. Meanwhile, differences in grooming were limited to simply a matter of hairstyle by cropping the photographs.

A perusal of the main study photographic array in Appendix 7 will show that the differences perceived by the participants were by no means extreme. Similar, seemingly unimportant, differences could easily have had their effect on previous studies. In particular, grooming, could well be imagined to systematically vary between different pre-existing groups of people, such as students exhibiting high or low Machiavellianism for instance, and might be particularly relevant to traits of sociability.

In the cases of age, grooming and picture quality, the inequality between photograph types seems to have occurred through the use of methods of control that were clearly too crude to deal with the fine distinctions drawn by the participants between different levels of these variables. The experimenter ratings were only intended to eliminate obvious differences between the photograph types, a goal resulting from an underestimation of the detail to which the participants would perceive gradations in these variables and the extent to which these fine distinctions would be associated with the main dependent variables. The experimenter ratings involved, at most, a five-point scale, and, in the case of age, judgments by decade, a lack of exactness that was aggravated by a procedure where non-criminal and sexual serial killer photographs were matched within a one-point tolerance limit for grooming and

picture quality. Having now established that apparently small group differences in age and grooming can confound ratings of criminality and the psychopathic traits, future attempts at control should be more vigorous.

Perceived age could be controlled more strenuously using group ratings in whole years, similar to the procedure for measurement used in the main study questionnaire. Control over grooming could be enhanced through standardising hairstyles, perhaps through donning wigs or caps, slicking back hair, or, when pre-existing photographs are used, by digitally concealing differences. Efforts could be made to retain information such as the hairline and the shape of ears, although there might be some loss of clarity regarding the shape of the head. Future research should continue to exclude environmental traces like tattoos, while make-up could be standardised by removal, where possible; a suggested manipulation that is supported by research reporting a Dorian Gray effect in women attributable to cosmetic enhancement (Zebrowitz, Collins, & Dutta, 1998).

The results relating to lighting and picture quality also indicate that superficial differences in the photographs themselves might have an influence on ratings. This is a danger that should be especially guarded against in studies using archival photographs from more than one source, a method that could introduce differences in photographic procedures. Where possible, photographs should be taken using the same procedures at a single location. Otherwise, selection of photographs from the results of group ratings would probably prove sufficient.

The difference in ratings of physical attractiveness that was found between the photograph types is not obviously explicable by inadequacies of method. The preliminary participants were older than the main study participants, but there is no indication that participant age, within the range evident in this study at least, substantially affects ratings of attractiveness (Berscheid & Walster, 1974; Sorrell & Nowak, 1981). Participant age is unlikely to have caused the contrast in ratings. There are, however, three explanations that appear more credible. First, the main study participants may simply have taken more care in their ratings, thus revealing more subtle differences; they had only 18 photographs to rate compared with 110 in the preliminary study. Second, the level of contrast between those being judged can

have an effect on ratings of physical attractiveness (Berscheid & Walster, 1974). A reduced range of attractiveness in the main study photographs may have encouraged ratings at a finer level of detail than in the preliminary study, allowing the indication of smaller differences in physical attractiveness between the sexual serial killers and non-criminals. Third, the tasks in the main study involving rating the photographs on the psychopathic traits may have led to the development of personality profiles for the photographs that flavoured ratings of physical attractiveness and caused an overall difference between the photograph types. This is supported by the finding that knowledge of social deviancy can affect ratings of attractiveness (Kowner, 1998).

If attributions of personality do affect ratings of physical attractiveness, the use of one group of participants to rate any variables to be controlled and a separate group to rate the main dependent variables would enhance similar studies in the future. The other two possible explanations for the differences in physical attractiveness suggest that the number of photographs rated in preliminary and main studies should probably be more equal in number, necessitating a greater number of preliminary participants, each rating a subset of the preliminary photographs.

An ancillary result from this study suggests a cause for concern regarding the interpretation of those pieces of research linking physical attractiveness to behavioural outcomes, for example those relating physical attractiveness to delinquency (Agnew, 1984; Cavior & Howard, 1973; Zebrowitz, Andreoletti et al., 1998; Zebrowitz & Lee, 1999). In the present investigation, there was a large positive correlation between ratings of expression and physical attractiveness. This was despite the experimenter's impression that the grand majority of the targets would probably have been described as having neutral expressions in those previous studies where expression was purportedly controlled but was not specifically measured. The implication is that ratings of physical attractiveness in other investigations may have been influenced by the participants' perceptions of emotional expressions, which could provide a feasible alternative explanation for any effects found. That is, personality could be a common causative factor for both the dependent variable (e.g. delinquency) and physical attractiveness, through the emotional expression of that personality. Such a possibility is supported by a study indicating that smiling raises ratings of attractiveness (Reis et al., 1990) and is also consistent with the finding that targets

who are asked to pose neutral expressions can nonetheless be judged to be displaying emotion (Malatesta et al., 1987).

In addition to traditional attempts to control expression through instructions or otherwise limiting obvious displays, it would appear prudent for future research to require participants to rate the expressions of the targets as a manipulation check, as was done in the present study. However, in order to ensure the wisdom of this course, the relationship between expression and physical attractiveness should first be experimentally confirmed. This might involve manipulating the emotions of targets who are instructed to keep a neutral expression and then measuring judges' perceptions of their emotions and attractiveness. The possibility that physical attractiveness causes misperception of emotional expression could also be explored through much the same procedure, by comparing judgments of emotion for targets of differing attractiveness instructed to keep their faces neutral while sharing the same induced mood.

Correlational Ambiguities

The cognitive theories of consensus in judgments based on the permanent features of the face (i.e. Nakdimen, 1984; Secord, 1958; Zebrowitz, 1997) share the understanding that qualities, such as those that were represented by the physical variables, initiate cognitions that bias physiognomic judgments. That is, according to the cognitive explanations for consensus, variations in the physical variables would *cause* variations in ratings of criminality and the possession of psychopathic traits. The data were consistent with the participants basing their ratings of psychopathic traits and criminality on qualities of the portraits that were measured by the physical variables. However, as the ratings were made simultaneously by the same people and there was no experimental control of the physical variables in the stimulus photographs, it is impossible to definitively conclude that the physical variables biased the participants' judgments.

As an alternative, it is quite possible that the correlations between the physical variables and ratings of psychopathic traits and criminality were the results of participants' judging certain photographs to represent criminals, psychopaths, or

simply "bad men" and then these conclusions affecting ratings on the physical variables. Certainly, earlier studies have been criticised because judges rated the targets on physical attractiveness after they were aware of their criminal histories (e.g. Agnew, 1984; Stewart, 1980, 1985), as knowledge of criminality might preclude perceptions of high attractiveness (Kowner, 1998). This direction of causation, while impossible to rule out, was attenuated in the present study by three aspects of the design. First, the participants were never informed that they were rating the photographs on traits relating to psychopathy. Being introductory students, there is no reason to suspect that they had knowledge regarding the PCL-R and none of them proffered any indication that he or she had connected the rating tasks to psychopathy. Second, the items relating to the PCL-R and the physical variables were randomly interspersed in the questionnaire and thus none of the physical or PCL-R based variables was systematically rated after its fellows. This meant that the physical variables were not uniformly rated after each participant had developed a full profile of the targets' personality, nor was the opposite ever true. It might even be hoped that the sheer number of ratings required and the random order of items served to discourage an appreciation of the theme behind the scales and somewhat reduced the likelihood that generally negative views of a target's personality shaded ratings of his appearance. Third, and most importantly, knowledge that some of the photographs portrayed people who had committed serious crimes was postponed until after all the other ratings had been made, and the participants were specifically requested not to refer back to their earlier answers. Having regard to these three design features, it is unlikely that perceptions of criminality or psychopathic traits commonly affected ratings of the physical variables, but some effect of general perceptions regarding the desirability of personality might well have had an influence.

The presence of the two different types of photograph, forming a known discontinuity in real criminality and psychopathy, provides a threat to the interpretation of the associations between the physical variables and psychopathic traits and criminality that is more specific to this study's individual design. It is true of all correlational designs that any associations discovered might be attributable to some underlying causative third variable. In this study, the data did not completely support physiognomic perception independent of the physical variables because certain physical variables that covaried with psychopathic traits and criminality were also

rated unequally across photograph type. However, it cannot be concluded from this that the participants were completely incapable of such physiognomic perception. The fact that the photographs included sexual serial killers and non-criminals and the participants may have been able to distinguish between them must be recognised as at least a possible cause of the relationships between ratings of psychopathic traits, criminality and the physical variables. For instance, the participants may have tended to underestimate the sexual serial killers' ages or levels of physical attractiveness because they could accurately perceive that they were antisocial, without reference to the physical variables, and also tended to rate antisocial people as younger and less attractive.

Nonetheless, some aspects of the data argue against the force of this possibility. First, as discussed above, group comparisons between the confused and correctly perceived photographs of different type were more consistent with a connection between *perceived* psychopathic traits or criminality and the physical variables. Second, only some of the physical variables were not balanced between the photograph types and the differences, while statistically significant, were not necessarily practically very large. Notably, there was only a three-year mean difference in perceived age between the sexual serial killers and non-criminals, compared with an overall range in the mean ages of all the photographs of 28 years. It seems unlikely that the effect sizes found for associations between age and psychopathic traits and criminality can be adequately explained by identifying photograph type as a common cause of both ratings of age and the main dependent variables.

To summarise, while the associations found may agree with what could be expected if the participants were making ratings of psychopathic traits and criminality which were biased by over-generalisations from physical qualities of the photographs and their subjects, the results can provide only qualified support for the cognitive theories of consensus. Still, this sense of qualification is, after all, shared in general by all correlational studies and the design does allow the identification of variables that might have particular relevance to perceptions of criminality and psychopathy. Such variables should be kept in mind as possible confounds in later research and might be worthy of further investigation themselves.

Separating Physiognomy from Self-fulfilling Prophecies

The data from this study were more consistent with accuracy resulting from the effects of over-generalisation and other cognitive biases than with physiognomy. However, as physiognomy could not be ruled out as an explanation for accuracy, there would be some purpose to its replication. Any such replication should include the stronger design features that have been detailed above, to ensure that the physical variables would be equalised across the photograph types and that the acts of rating the physical variables and the main dependent variables would be independent.

Apart from an attempt to replicate this study, there may be other ways to separate physiognomy from accuracy stemming from over-generalisation. The most obvious test of structural physiognomy would involve facial judgments of extremely young children, who could not have been exposed to the effects of self-fulfilling prophecies. Only one study (Kessen, 1957) has investigated the perception of personality for children (aged 14 to 23 months) and, while it used nude full-body photographs and was thus not restricted to facial assessment, accurate perception was supported. Perhaps temperament in even younger children might be amenable to physiognomic perception. Temperament would be relevant to affordances that impact on the parent-child relationship and therefore should be related to survival and accurately perceived according to an ecological approach to person perception. In this proposed investigation the physical variables that were indicated as possible confounds in the present study would be largely irrelevant. Only the possibility of perceiving transient emotional expression would need to be controlled, easily accomplished by taken photographs of peacefully sleeping children or, perhaps, deceased or anaesthetised children who would not be experiencing emotions.

Expressional physiognomy would probably be best demonstrated in an extended replication of the study by Malatesta et al. (1987) that specifically controlled the possible confounds of age, grooming, physical attractiveness, babyfaceness, transient emotional expression and biological causation. As suggested above, transient emotional expression could be controlled by taking photographs of the targets during a non-conscious state and grooming could be regulated by obscuring hairstyle. Age, physical attractiveness, babyfaceness and a shared biological cause for face and

personality could all be controlled by using photographs of monozygotic twins, with equivalency of attractiveness and babyfaceness checked by ratings provided by a separate sample of participants (to ensure independence from perceptions of character).

A shift in emphasis might be warranted, from trying to determine whether people can perceive underlying character through physiognomy toward finding physical signs of personality. The most intriguing and least equivocal data linking fixed facial appearance to personality have arisen from studies using precise measurements of structural features of the face, for example jaw structure (McCabe, 1928; Squier & Mew, 1981) or the existence of minor physical anomalies (e.g. Firestone & Peters, 1983; Halverson & Victor, 1976; Kandel et al., 1989; Paulhus & Martin, 1986; Waldrop et al., 1978; Waldrop et al., 1968). Whether or not people use these links in reading personality from the face, they appear to be worthy of study; such physical features provide a refreshing opportunity for accurate measurement that is denied to researchers dealing with global assessments of the face. Despite its turgid history, a return to the methods of traditional structural physiognomy might prove complementary to those investigations using global facial descriptions. Both approaches might connect facial appearance to personality or behaviour, but the exact measurements used in traditional structural physiognomy may yet uncover a biological association that is a more reliable predictor of personality in the individual than the real trait-appearance correlations attributable to general facial descriptions and self-fulfilling prophecies.

Conclusion

This study was predicated upon the belief that, without separating the effects of contemporaneous participant bias and self-fulfilling prophecies from the facial perception of character, there can be no real confidence in findings that purportedly support physiognomic accuracy. The results of this study demonstrated accuracy in facial judgments but did not provide unqualified support for physiognomic perceptions of personality and behaviour that were independent of those physical characteristics of the stimuli that seem likely to elicit non-physiognomic biases.

Patterns in the data were entirely consistent with perceptions of character being linked to physical variables of the photographs and their subjects, independent of the actual characters of the people portrayed. They were also consistent with over-generalisation from adaptive perceptions relating to age, health and emotional expression. In essence, the validity of physiognomic perception is called into question. This question will only be answered through future research that adequately controls the effects of biases relating to global descriptions such as physical attractiveness and non-physiognomic confounds like grooming and emotional expression.

It may well be that any accuracy that has been found in studies using facial photographs is best explained by interactions between contemporaneous bias or probabilistic knowledge and uncontrolled confounds, or through the self-fulfilling effects of facial stereotypes. If this is true, then, within the context of an ecological approach to person perception, the continued efforts of the populace to read character directly from the face are probably maladaptive consequences of an attunement to features involved in other adaptive perceptions of health, fitness, age and so on (Zebrowitz, 1997). Presumably, any direct perception of personality would be dependent upon more dynamic sources of information than the features of static faces.

References

Adams, G. R. (1977). Physical attractiveness research: Toward a developmental social psychology of beauty. Human Development, 20, 217-239.

Agnew, R. (1984). Appearance and delinquency. Criminology: An Interdisciplinary Journal, 22, 421-440.

Alley, T. R. (1988). Physiognomy and social perception. In T. R. Alley (Ed.), Social and applied aspects of perceiving faces (pp. 167-186). Hillsdale, NJ: Lawrence Erlbaum Associates.

Allport, G. W. (1937). Personality: A psychological interpretation. New York: Henry Holt and Company.

Alterman, A. I., Cacciola, J. S., & Rutherford, M. J. (1993). Reliability of the revised psychopathy checklist in substance abuse patients. Psychological Assessment, 5, 442-448.

American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: American Psychiatric Association.

Atwan, R. (1984). Physiognomy, photography, and prostitution: Cesare Lombroso and the female offender. Research Communications in Psychology, Psychiatry, and Behavior, 9, 353-364.

Berry, D. S. (1990). Taking people at face value: Evidence for the kernal of truth hypothesis. Social Cognition, 8, 343-361.

Berry, D. S. (1991). Accuracy in social perception: Contributions of facial and vocal information. Journal of Personality and Social Psychology, 61, 298-307.

Berry, D. S., & Brownlow, S. (1989). Were the physiognomists right?: Personality correlates of facial babyishness. Personality and Social Psychology Bulletin, 15, 266-279.

Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces: The impact of age-related craniofacial changes on social perception. Psychological Bulletin, 100, 3-18.

Berry, D. S., & Wero, J. L. F. (1993). Accuracy in face perception: A view from ecological psychology. Journal of Personality, 61, 497-520.

Berry, D. S., & Zebrowitz-McArthur, L. (1988a). The impact of age-related craniofacial changes on social perception. In T. R. Alley (Ed.), Social and applied aspects of perceiving faces.(pp. 63-87). Hillsdale, NJ: Lawrence Erlbaum Associates.

Berry, D. S., & Zebrowitz-McArthur, L. (1988b). What's in a face?: Facial maturity and the attribution of legal responsibility. Personality and Social Psychology Bulletin, 14, 23-33.

Berscheid, E., & Walster, E. (1974). Physical attractiveness. In L. Berkowitz (Ed.), Advances in social psychology (Vol. 7, pp. 157-215). New York: Academic Press.

Black, J. A., Lane, B., & Pender, P. (1994a). The corpse collector. Real-Life Crimes...and how they were solved, 6, 1945-1953.

Black, J. A., Lane, B., & Pender, P. (1994b). Who killed Helen Smith?. Real-Life Crimes...and how they were solved, 6, 1958-1966.

Bond, C. F., Berry, D. S., & Omar, A. (1994). The kernal of truth in judgments of deceptiveness. Basic and Applied Social Psychology, 15, 523-534.

Boor, M. (1976). Beautiful is not dangerous, beauty is not talent: Two failures to replicate physical attractiveness effects. JSAS: Catalog of Selected Documents in Psychology, 6, 109.

Borod, J. C. (1993). Cerebral mechanisms underlying facial, prosodic, and lexical emotional expression: A review of neuropsychological studies and methodological issues. Neuropsychology, 7, 445-463.

Brandenburg, G. C. (1926). Do physical traits portray character? Industrial Psychology, 1, 580-588.

Brandt, A. (1980). Face reading : The persistence of physiognomy. Psychology Today, 14, 90-96.

Bray, R. M., & Kerr, N. L. (1982). Methodological considerations in the study of the psychology of the courtroom. In N. L. Kerr & R. M. Bray (Eds.), The psychology of the courtroom (pp. 267-323). New York: Academic Press.

Bull, R. (1982). Physical appearance and criminality. Current Psychological Reviews, 2, 269-282.

Bull, R. H. C., & Green, J. (1980). The relationship between physical appearance and criminality. Medicine, Science and the Law, 20, 79-83.

Bull, R., & Rumsey, N. (1988). The social psychology of facial appearance. New York: Springer-Verlag.

Cash, T. F., Kehr, J. A., Polyson, J., & Freeman, V. (1977). Role of physical attractiveness in peer attribution of psychological disturbance. Journal of Consulting and Clinical Psychology, 6, 987-993.

Cavior, N., & Howard, L. R. (1973). Facial attractiveness and juvenile delinquency among black and white offenders. Journal of Abnormal Child Psychology, 1, 202-213.

Cherulnik, P. D., Turns, L. C., & Wilderman, S. K. (1990). Physical appearance and leadership:Exploring the role of apearance-based attribution in leader emergence. Journal of Applied Social Psychology, 20, 1530-1539.

Cherulnik, P. D., Way, J. H., Ames, S., & Hutto, D. B. (1981). Impressions of high and low Machiavellian men. Journal of Personality, 49, 388-400.

Cleckley, H. (1964). The mask of sanity (4th Edition). Saint Louis: CV Mosby Co.

Cleeton, G. U., & Knight, F. B. (1924). Validity of character judgments based on external criteria. Journal of Applied Psychology, 8, 215-231.

Cook, S. W. (1939). The judgment of intelligence from photographs. Journal of Abnormal and Social Psychology, 34, 384-389.

Cooke, D. J., & Michie, C. (1997). An item response theory analysis of the Hare psychopathy checklist-revised. Psychological Assessment, 9, 3-14.

Cornell, D. G., Warren, J., Hawk, G., Stafford, E., Oram, G., & Pine, D. (1996). Psychopathy in instrumental and reactive violent offenders. Journal of Consulting and Clinical Psychology, 64, 783-790.

Darby, B. W., & Jeffers, D. (1988). The effects of defendant and juror attractiveness on simulated courtroom trial decisions. Social Behavior and Personality, 16, 39-50.

Deitz, S. R., & Byrnes, L. E. (1981). Attribution of responsibility for sexual assault: The influence of observer empathy and defendant occupation and attractiveness. The Journal of Psychology, 108, 17-29.

DeSantis, A., & Kayson, W. A. (1997). Defendants' characteristics of attractiveness, race, and sex and sentencing decisions. Psychological Reports, 81, 679-683.

Dickey-Bryant, L., Lautenschlager, G. J., Mendoza, J. L., & Abrahams, N. (1986). Facial attractiveness and its relation to occupational success. Journal of Applied Psychology, 71, 16-19.

Dion, K. K. (1973). Young children's stereotyping of facial attractiveness. Developmental Psychology, 9, 183-188.

Dolan, M. (1994). Psychopathy- a neurobiological perspective. British Journal of Psychiatry, 165, 151-159.

Douglas, J., & Olshaker, M. (1996). Mindhunter. London: Heinemann.

Dunkle, J. H., & Francis, P. L. (1996). "Physical attractiveness stereotype" and the attribution of homosexuality revisited. Journal of Homosexuality, 30, 13-29.

Efran, M. G. (1974). The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. Journal of Research in Personality, 8, 45-54.

Eisenstaedt, A. (1976). Eisenstaedt's album: Fifty years of friends and acquaintances. London: Thames & Hudson.

Eisenstaedt, A. (1980). Witness to our time. London: Secker & Warburg.

Farina, A., Austad, C., Burns, G. L., Bugglin, C., & Fischer, E. H. (1986). The role of physical attractiveness in the readjustment of discharged psychiatric patients. Journal of Abnormal Psychology, 95, 139-143.

Feingold, A. (1992). Good-looking people are not what we think. Psychological Bulletin, 111, 304-341.

Fink, L. (1984). Social graces: Photographs by Larry Fink. New York: Aperture.

Firestone, P., & Peters, S. (1983). Minor physical anomalies and behavior in children: A review. Journal of Autism and Developmental Disorders, 13, 411-425.

Forth, A. E., Hart, S. D., & Hare, R. D. (1990). Assessment of psychopathy in male young offenders. Journal of Consulting and Clinical Psychology, 3, 342-344.

Fox, J. A., & Levin, J. (1994). Overkill: Mass murder and serial killing exposed. New York: Plenum Press.

Ganey, T. (1989). St. Joseph's children: A true story of terror and justice. New York: Carol Publishing Group.

Geberth, V. J., & Turco, R. N. (1997). Antisocial personality disorder, malignant narcissism, and serial murder. Journal of Forensic Sciences, 42, 49-60.

Gekoski, A. (1998). Murder by numbers: British serial sex killers since 1950: Their childhoods, their lives, their crimes. London: Andre Deutsch.

Gerdes, E. P., Dammann, E. J., & Helig, K. E. (1988). Perceptions of rape victims and assailants: Effects of physical attractiveness, acquaintance, and subject gender. Sex Roles, 19, 141-153.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston: Houghton Mifflin.

Goldstein, A. G., Chance, J. E., & Gilbert, B. (1984). Facial stereotypes of good guys and bad guys: A replication and extension. Bulletin of the Psychonomic Society, 22, 549-552.

Greenwood, J. D. (1983). Role-playing as an experimental strategy in social psychology. European Journal of Social Psychology, 13, 235-254.

Gurnee, H. (1934). An analysis of the perception of intelligence. in the face. Journal of Social Psychology, 5, 82-89.

Hall, A. (1993). Murder and madness. Enderby, UK: Blitz Editions.

Halverson, C. F., & Victor, J. B. (1976). Minor physical anomalies and problem behavior in elementary school children. Child Development, 47, 281-285.

Hare, R. D. (1991). The Hare psychopathy checklist-revised (PCL-R). Toronto, Ontario: Multi-Health Systems.

Hare, R. D. (1996). Psychopathy: A clinical construct whose time has come. Criminal Justice and Behavior, 23, 25-54.

Hare, R. D. (1998). Psychopathy, affect and behavior. In D. J. Cooke, A. E. Forth, & R. D. Hare (Eds.), Psychopathy: Theory, research and implications for society (pp. 355-373). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Hare, R. D., Harpur, T. J., Hakstian, A. R., Forth, A. E., Hart, S. D., & Newman, J. P. (1990). The revised psychopathy checklist: Reliability and factor structure. Journal of Consulting and Clinical Psychology, 2, 338-341.

Harpur, T. J., & Hare, R. D. (1990). Psychopathy and attention. In J. T. Enns (Ed.), The development of attention: Research and theory (pp. 429-444).

Harpur, T. J., & Hare, R. D. (1994). Assessment of psychopathy as a function of age. Journal of Abnormal Psychology, 103, 604-609.

Harpur, T. J., Hare, R. D., & Hakstian, A. R. (1989). Two-factor conceptualization of psychopathy: Construct validity and assesssment implications. Journal of Consulting and Clinical Psychology, 1, 6-17.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1994). Psychopathy as a taxon: Evidence that psychopaths are a discrete class. Journal of Consulting and Clinical Psychology, 62, 387-397.

Hart, J. (1995). 50 Portrait lighting techniques for pictures that sell (rev. ed.). New York: Watson-Guptill Publications.

Hart, S. D. (1998). Psychopathy and risk for violence. In D. J. Cooke, A. E. Forth, & R. D. Hare (Eds.), Psychopathy: Theory, research and implications for society (pp. 355-373). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Hart, S. D., Forth, A. E., & Hare, R. D. (1990). Performance of criminal psychopaths on selected neuropsychological tests. Journal of Abnormal Psychology, 99, 374-379.

Hart, S. D., & Hare, R. D. (1989). Discriminant validity of the psychopathy checklist in a forensic psychiatric population. Journal of Consulting and Clinical Psychology, 1, 211-218.

Hart, S. D., & Hare, R. D. (1997). Psychopathy: Assessment and association with criminal conduct. In E. M. Stoff, J. Breiling & J. D. Maser (Eds.), Handbook of antisocial behavior (pp. 22-35). New York: John Wiley & Sons, Inc.

Hart, S. D., Kropp, P. R., & Hare, R. D. (1988). Performance of male psychopaths following conditional release from prison. Journal of Consulting and Clinical Psychology, 56, 227-232.

Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. Legal and Criminological Psychology, 3, 139-170.

Hemphill, J. F., Templeman, R., Wong, S., & Hare, R. D. (1998). Psychopathy and crime: Recidivism and criminal careers. In D. J. Cooke, A. E. Forth, & R. D. Hare (Eds.), Psychopathy: Theory, research and implications for society (pp. 375-399). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Hochberg, J., & Galper, R. E. (1974). Attribution of intention as a function of physiognomy. Memory and Cognition, 2, 39-42.

Howland, E. W., Kosson, D. S., Patterson, C. M., & Newman, J. P. (1993). Altering a dominant response: Performance of psychopaths and low-socialization college students on a cued reaction time task. Journal of Abnormal Psychology, 102, 379-387.

Jackson, R. (1995). True crimes. Leichhardt, Australia: Quintet Publishing.

Jacobson, M. B. (1981). Effects of victim's and defendant's physical attractiveness on subjects' judgments in a rape case. Sex Roles, 7, 247-255.

Jessel, D., Blake, S., Morton, J., & Waddell, B. (1989a). Look inside murder casebook: Investigations into the ultimate crime. London: Marshall Cavendish Partworks.

Jessel, D., Blake, S., Morton, J., & Waddell, B. (1989b). The Yorkshire ripper. Murder Casebook: Investigations into the Ultimate Crime, 1, 4-35.

Jessel, D., Blake, S., Morton, J., & Waddell, B. (1990a). The Boston strangler. Murder Casebook: Investigations into the Ultimate Crime, 1, 148-179.

Jessel, D., Blake, S., Morton, J., & Waddell, B. (1990b). Crimes of Sadism. Murder Casebook: Investigations into the Ultimate Crime, 1, 508-539.

Jessel, D., Blake, S., Morton, J., & Waddell, B. (1990c). Dennis Nilsen. Murder Casebook: Investigations into the Ultimate Crime, 1, 40-71.

Jessel, D., Blake, S., Morton, J., & Waddell, B. (1990d). John Reginald Christie. Murder Casebook: Investigations into the Ultimate Crime, 1, 112-143.

Jessel, D., Wilson, C., Morton, J., & Waddell, B. (1990a). The Hillside stranglers. Murder Casebook: Investigations into the Ultimate Crime, 2, 616-647.

Jessel, D., Wilson, C., Morton, J., & Waddell, B. (1990b). Madness for two. Murder Casebook: Investigations into the Ultimate Crime, 2, 1264-1295.

Jessel, D., Wilson, C., Morton, J., & Waddell, B. (1990c). Peter Manuel. Murder Casebook: Investigations into the Ultimate Crime, 2, 904-935.

Jessel, D., Wilson, C., Morton, J., & Waddell, B. (1991a). Bodies in the undergrowth. Murder Casebook: Investigations into the Ultimate Crime, 6, 2920-2951.

Jessel, D., Wilson, C., Morton, J., & Waddell, B. (1991b). Deadly games. Murder Casebook: Investigations into the Ultimate Crime, 5, 2632-2663.

Jessel, D., Wilson, C., Morton, J., & Waddell, B. (1991c). Eric Edgar Cooke. Murder Casebook: Investigations into the Ultimate Crime, 6, 3208-3239.

Jessel, D., Wilson, C., Morton, J., & Waddell, B. (1991d). Highway of death. Murder Casebook: Investigations into the Ultimate Crime, 6, 2848-2879.

Jessel, D., Wilson, C., Morton, J., & Waddell, B. (1991e). William Heirens. Murder Casebook: Investigations into the Ultimate Crime, 5, 2236-2267.

Jones, W. H., Hansson, R. O., & Phillips, A. L. (1978). Physical attractiveness and judgments of psychopathology. Journal of Social Psychology, 105, 79-84.

Kandel, E., Brennan, P. A., Mednick, S. A., & Michelson, N. M. (1989). Minor physical anomalies and recidivistic adult violent criminal behavior. Acta Psychiatrica Scandinavica, 79, 103-107.

Keating, C. F. (1985). Gender and the physiognomy of dominance and attractiveness. Social Psychology Quarterly, 48, 61-70.

Kessen, W. (1957). The role of experience in judging children's photographs. Journal of Abnormal and Social Psychology, 54, 375-379.

Knight, R. (1932). Character and the face. Human Factor, 6, 413-422.

Kosson, D. S., & Newman, J. P. (1986). Psychopathy and the allocation of attentional capacity in a divided-attention situation. Journal of Abnormal Psychology, 95, 257-263.

Kowner, R. (1998). Effects of social deviance labels on judgments of facial attractiveness: A comparison of labelling procedures using Japanese raters. International Journal of Psychology, 33, 1-16.

Landy, D., & Aronson, E. (1969). The influence of the character of the criminal and victim on the decisions of simulated jurors. Journal of Experimental Social Psychology, 5, 141-152.

Lane, B., & Gregg, W. (1996). The new encyclopedia of serial killers. London: Headline.

Langlois, J. H. (1986). From the eye of the beholder to behavioral reality: Development of social behaviors and social relations as a function of physical attractiveness. In C. P. Herman, M. P. Zanna, & E. T. Higgins (Eds.), Physical appearance, stigma and social behavior: The Ontario Symposium (Vol. 3, pp. 23-51). Hillsdale, NJ: Erlbaum.

Langlois, J. H., & Downs, A. C. (1979). Peer relations as a function of physical attractiveness: The eye of the beholder or behavioral reality? Child Development, 50, 409-418.

Langlois, J. H., Roggman, L. A., Casey, R. J., Ritter, J. M., Rieser-Danner, L. A., & Jenkins, V. Y. (1987). Infant preferences for attractive faces: Rudiments of a stereotype? Developmental Psychology, 23, 363-369.

Langlois, J. H., & Stephan, C. W. (1981). Beauty and the beast: The role of physical attractiveness in the development of peer relations and social behavior. In S. S. Brehm, S. M. Kassim, & F. X. Gibbons (Eds.), Developmental social psychology (pp. 152-168). New York: Oxford University Press.

LaPierre, D., Braun, C. M. J., & Hodgins, S. (1995). Ventral frontal deficits in psychopathy: Neuropsychological test findings. Neuropsychologia, 33, 139-151.

Laser, P. S., & Mathie, V. A. (1982). Face facts: An unbidden role for features in communication. Journal of Nonverbal Behavior, 7, 3-19.

Lasswell, T. E., & Parshall, P. F. (1962). The perception of social class from photographs. Sociology and Social Research, 45, 407-414.

Lerner, R. M., & Lerner, J. V. (1977). Effects of age, sex, and physical attractiveness on child-peer relations, academic performance, and elementary school adjustment. Developmental Psychology, 13, 585-590.

Leventhal, G., & Krate, R. (1977). Physical attractiveness and severity of sentencing. Psychological Reports, 40, 315-318.

Levin, J., & Fox, J. A. (1986). Mass murder: America's growing menace. New York: Plenum Press.

Lilienfeld, S. O. (1994). Conceptual problems in the assessment of psychology. Clinical Psychology Review, 14, 17-38.

Lilienfeld, S. O. (1998). Methodological advances and developments in the assessment of psychopathy. Behaviour Research and Therapy, 36, 99-125.

Lynam, D. R. (1996). Early identification of chronic offenders: Who is the fledgling psychopath? Psychological Bulletin, 120, 209-234.

McArthur, L. Z. (1982). Judging a book by its cover: A cognitive analysis of the relationship between physical appearance and stereotyping. In A. H. Hastorf & A. M. Isen (Eds.), Cognitive social psychology (pp. 149-211). New York: Elsvier/North-Holland.

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. Psychological Review, 90, 215-238.

McArthur, L. Z., & Berry, D. S. (1987). Cross-cultural agreement in perceptions of babyfaced adults. Journal of Cross-Cultural Psychology, 18, 165-192.

McCabe, F. E. (1928). Anatomical and other alleged signs of aptitude. In C. L. Hull (Ed.), Aptitude testing (pp. 111-155). New York: World Book.

McCord, W. M., & McCord, J. (1956). Psychopathy and delinquency. New York: Grune & Stratton.

MacCoun, R. J. (1990). The emergence of extralegal bias during jury deliberation. Criminal Justice and Behavior, 17, 303-314.

McGee, A-M., & Skinner, M. (1987). Facial asymmetry and the attribution of personality traits. British Journal of Social Psychology, 26, 181-184.

McHoskey, J. W., Worzel, W., & Szyarto, C. (1998). Machiavellianism and psychopathy. Journal of Personality and Social Psychology, 74, 192-210.

Macrae, C. N., & Shepherd, J. W. (1989). Do criminal stereotypes mediate juridic judgments? British Journal of Social Psychology, 28, 189-191.

Mainwaring, M. (1980). 'Phys/phren'-why not to take each other at face value. Smithsonian, 11, 193-212.

Malatesta, C. Z., Fiore, M. J., & Messina, J. J. (1987). Affect, personality, and facial expressive characteristics of older people. Psychology and Aging, 2, 64-69.

Mason, D. J. (1959). Judgments of leadership based upon physiognomic cues. Journal of Abnormal and Social Psychology, 54, 273-274.

Mazzella, R., & Feingold, A. (1994). The effects of physical attractiveness, race, socioeconomic status, and gender of defendants and victims on judgments of mock jurors: A meta-analysis. Journal of Applied Social Psychology, 24, 1315-1344.

Mazur, A., Mazur, J., & Keating, C. (1984). Military rank attainment of a West Point class: Effects of cadets' physical features. American Journal of Sociology, 90, 125-150.

Michals, D. (1988). Album: The portraits of Duane Michals 1958-1988. Pasadena, CA: Twelvetrees Press.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. Psychological Review, 100, 674-701.

Mones, P. (1995). Stalking justice. New York: Pocket Books.

Mook, D. G. (1983). In defense of external validity. American Psychologist, 38, 379-413.

Moore, C. H., Wuensch, K. L., Hedges, R. M., & Castellow, W. A. (1994). The effects of physical attractiveness and social desirability on judgments regarding a sexual harassment case. Journal of Social Behavior and Personality, 9, 715-730.

Morath, I. (1986). Portraits. New York: Aperture.

Morley, L. (1989). Lewis Morley: photographer of the sixties. London: National Portrait Gallery Publications.

Nakdimen, K. A. (1984). The physiognomic basis of sexual stereotyping. American Journal of Psychiatry, 141, 499-503.

Nash, H. (1958). Stereotyped associations to schematic faces. Journal of Genetic Psychology, 93, 149-152.

Newman, J. P. (1998). Psychopathic behavior: An information processing perspective. In D. J. Cooke, A. E. Forth, & R. D. Hare (Eds.), Psychopathy: Theory,

research and implications for society (pp. 355-373). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Newman, J. P., & Kosson, D. S. (1986). Passive avoidance learning in psychopathic and nonpsychopathic offenders. Journal of Abnormal Psychology, 95, 252-256.

Newman, J. P., Patterson, C. M., & Kosson, D. S. (1987). Response perseveration in psychopaths. Journal of Abnormal Psychology, 96, 145-148.

Newman, J. P., Schmitt, W. A., & Voss, W. D. (1997). The impact of motivationally neutral cues on psychopathic individuals: Assessing the generality of the response modulation hypothesis. Journal of Abnormal Psychology, 106, 563-575.

Patrick, C. J., Bradley, M. M., & Lang, P. J. (1993). Emotion in the criminal psychopath: Startle reflex modulation. Journal of Abnormal Psychology, 102, 82-92.

Patrick, C. J., Cuthbert, B. N., & Lang, P. J. (1994). Emotion in the criminal psychopath: Fear image processing. Journal of Abnormal psychology, 103, 523-534.

Patrick, C. J., Zempolich, K. A., & Levenston, G. K. (1997). Emotionality and violent behavior in psychopaths. In A. Raine, P. A. Brennan, D. P. Farrington & S. A. Mednick (Eds.), Biosocial bases of violence (pp. 145-161). New York: Plenum Press.

Paulhus, D. L., & Martin, C. L. (1986). Predicting adult temperament from minor physical anomalies. Journal of Personality and Social Psychology, 50, 1235-1239.

Pender, P., Innes, B., & Lane, B. (1993a). Donald Neilson: The black panther. Real-Life Crimes...and how they were solved, 1, 53-62.

Pender, P., Innes, B., & Lane, B. (1993b). Harvey Glatman: Glamour girl slayer. Real-Life Crimes...and how they were solved, 1, 72-74.

Pender, P., Smyth, F., & Meenan, M. (1994a). The gay slayer. Real-Life Crimes...and how they were solved, 7, 2253-2261.

Pender, P., Smyth, F., & Meenan, M. (1994b). Murder in the family. Real-Life Crimes...and how they were solved, 7, 2262-2269.

Premoli, A. (Ed.) (1997). L'Uomo Vogue, July 1997.

Raine, A, Brennan, P., & Farrington, D. P. (1997). Biosocial bases of violence: Conceptual and theoretical issues. In A. Raine, P. A. Brennan, D. P. Farrington & S. A. Mednick (Eds.), Biosocial bases of violence (pp. 145-161). New York: Plenum Press.

Raine, A., O'Brien, M., Smiley, N., Scerbo, A., & Chan, C-J. (1990). Reduced lateralization in verbal dichotic listening in adolescent psychopaths. Journal of Abnormal Psychology, 99, 272-277.

Ramsey, C. (Ed.) (1999). Hair (British Edition), Dec/Jan 1999. Haywards Heath, UK: IPC Magazines.

Ray, W. S. (1958). Judgments of intelligence based on brief observations of physiognomy. Psychological Reports, 4, 478.

Reis, H. T., Wilson, I. M., Monestere, C., Bernstein, S., Clark, K., Seidl, E., Franco, M., Gioioso, E., Freeman, L., & Radoane, K. (1990). What is smiling is beautiful and good. European Journal of Social Psychology, 20, 259-267.

Ressler, R. K., Burgess, A. W., Hartman, C. R., Douglas, J. E., & McCormack, A. (1986). Murderers who rape and mutilate. Journal of Interpersonal Violence, 1, 273-287.

Rhodes, G., & Lynskey, M. (1990). Face perception: Attributions, asymmetries and stereotypes. British Journal of Social Psychology, 29, 375-377.

Robertson, J. B. (Ed) (1992). Adams on criminal law. Wellington, NZ: Brooker & Friend.

Roll, S., & Verinis, J. S. (1971). Stereotypes of scalp and facial hair as measured by the semantic differential. Psychological Reports, 28, 975-980.

Saladin, M., Saper, Z., & Breen, L. (1988). Perceived attractiveness and attributions of criminality: What is beautiful is not criminal. Canadian Journal of Criminology, 30, 251-259.

Samuels, M. R. (1939). Judgments of faces. Character and Personality, 8, 18-27.

Sander, A. (1971). Men without masks: Faces of Germany 1910-1938. Greenwich, Conneticut: New York Graphic Society.

Sappenfield, B. R. (1977). Perception of openness, trustworthiness, and other characteristics in facial photographs. Perceptual and Motor Skills, 45, 195-200.

Scerbo, A., Raine, A., O'Brien, M., Chan, C-J., Rhee, C., & Smiley, N. (1990). Reward dominance and passive avoidance learning in adolescent psychopaths. Journal of Abnormal Child Psychology, 18, 451-463.

Secord, P. F. (1958). Facial features and inference processes in interpersonal perception. In R Tagiuri & L. Petrullo (Eds.), Person perception and interpersonal behavior (pp. 300-315). Stanford: Stanford University Press.

Secord, P. F., & Bevan, W. (1956). Personalities in faces: III. A cross-cultural comparison of impressions of physiognomy and personality in faces. Journal of Social Psychology, 43, 283-288.

Secord, P. F., Bevan, W., & Dukes, W. F. (1953). Occupational and physiognomic stereotypes in the perception of photographs. Journal of Social Psychology, 37, 261-270.

Secord, P. F., Dukes, W. F., & Bevin, W. (1954). Personalities in faces: I. An experiment in social perceiving. Genetic Psychology Monographs, 49, 231-279.

Secord, P. F., & Muthard, J. E. (1955a). Individual differences in the perception of women's faces. Journal of Abnormal and Social Psychology, 50, 238-242.

Secord, P. F., & Muthard, J. E. (1955b). Personalities in faces: IV. A descriptive analysis of the perception of women's faces and the identification of some physiognomic determinants. Journal of Psychology, 39, 269-278.

Secord, P. F., Stritch, T. F., & Johnson, L. (1960). The role of metaphorical generalization and congruency in the perception of facial characteristics. Journal of Social Psychology, 52, 329-337.

Serin, R. C. (1991). Psychopathy and violence in criminals. Journal of Interpersonal Violence, 6, 423-431.

Serin, R. C. (1992). The clinical application of the psychopathy checklist-revised (PCL-R) in a prison population. Journal of Clinical Psychology, 48, 637-641.

Serin, R. C., Peters, R. DeV., & Barbaree, H. E. (1990). Predictors of psychopathy and release outcome in a criminal population. Journal of Consulting and Clinical Psychology, 2, 419-422.

Shoemaker, D. J., South, D. R., & Lowe, J. (1973). Facial stereotypes of deviants and judgments of guilt or innocence. Social Forces, 51, 427-433.

Sigall, H., & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. Journal of Personality and Social Psychology, 31, 410-414.

Snowdon (1983). Sittings, 1979-1983. London: Weidenfeld & Nicolson.

Snyder, M. (1984). When belief creates reality. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 18, pp. 247-305). New York: Academic Press.

Sobieszek, R. A., & Irmas, D. (1994). The camera i: Photographic self-portraits from the Audrey & Sydney Irmas collection. New York: Harry N. Abrams.

Sorell, G. T., & Nowak, C. A. (1981). The role of physical attractiveness as a contributor to individual development. In R. M. Lerner & N. S. Busch-Rossnagel (Eds.), Individuals as producers of their development: A life span perspective (pp. 389-446). San Diego, CA: Academic Press.

Squier, R. W., & Mew, J. R. C. (1981). The relationship between facial structure and personality characteristics. British Journal of Social Psychology, 20, 151-160.

Stewart, J. E. (1980). Defendant's attractiveness as a factor in the outcome of criminal trials: An observational study. Journal of Applied Social Psychology, 10, 348-361.

Stewart, J. E. (1985). Appearance and punishment: The attraction-leniency effect in the courtroom. The Journal of Social Psychology, 125, 373-378.

Thompson, K. M. (1990). Refacing inmates: A critical appraisal of plastic surgery programs in prison. Criminal Justice and Behavior, 17, 448-466.

Thornton, G. R. (1939). The ability to judge crimes from photographs of criminals: A contribution to technique. Journal of Abnormal and Social Psychology, 34, 378-383.

Thornton, G. R. (1943). The effect upon judgments of personality traits of varying a single factor in a photograph. Journal of Social Psychology, 18, 127-148.

Topping, P., & Ritchie, J. (1989). Topping: The autobiography of the police chief in the moors murder case. London: Angus & Robertson Publishers.

Unger, R. K., Hilderbrand, M., & Madar, T. (1982). Physical attractiveness and assumptions about social deviance: Some sex-by-sex comparisons. Personality and Social Psychology Bulletin, 8, 293-301.

Waldrop, M. F., Bell, R. Q., McLaughlin, B., & Halverson, C. F. (1978). Newborn minor physical anomalies predict short attention span, peer aggression, and impulsivity at age 3. Science, 199, 563-565.

Waldrop, M. F., Pedersen, F. A., & Bell, R. Q. (1968). Minor physical anomalies and behavior in preschool children. Child Development, 39, 391-400.

Warren, J. I., Hazelwood, R. R., & Dietz, P. E. (1996). The sexually sadistic serial killer. Journal of Forensic Sciences, 41, 970-974.

West, A. M., & Hill, V. (1995). Out of the shadows. London: Simon & Schuster.

Wuensch, K. L., Castellow, W. A., & Moore, C. H. (1991). Effects of defendant attractiveness and type of crime on juridic judgment. Journal of Social Behavior and Personality, 6, 713-724.

Wuensch, K. L., Chia, R. S., Castellow, W. A., Chuang, C-J., & Cheng, B-S. (1993). Effects of physical attractiveness, sex, and type of crime on mock juror decisions: A replication with Chinese students. Journal of Cross-Cultural Psychology, 24, 414-427.

Yarmey, A. D. (1982). Eyewitness identification and stereotypes of criminals. In A. Trankell (Ed.), Reconstructing the past: the role of psychologists in criminal trials (pp. 205-225). Deventer, The Netherlands: Kluwer.

Yarmey, A. D. (1993). Stereotypes and recognition memory for faces and voices of good guys and bad guys. Applied Cognitive Psychology, 7, 419-431.

Zebrowitz, L. A. (1997). Reading faces: Window to the soul? Boulder, CO: Westview Press.

Zebrowitz, L. A., Andreoletti, C., Collins, M. A., Lee, S. Y., & Blumenthal, J. (1998). Bright, bad, babyfaced boys: Appearance stereotypes do not always yield self-fulfilling prophecy effects. Journal of Personality and Social Psychology, 75, 1300-1320.

Zebrowitz, L. A., & Collins, M. A. (1997). Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. Personality and Social Psychology Review, 1, 204-223.

Zebrowitz, L. A., Collins, M. A., & Dutta, R. (1998). The relationship between appearance and personality across the life span. Personality and Social Psychology Bulletin, 24, 736-749.

Zebrowitz, L., & Lee, S. Y. (1999). Appearance, stereotype-incongruent behavior, and social relationships. Personality and Social Psychology Bulletin, 25, 569-584.

Zebrowitz, L. A., Olson, K., & Hoffman, K. (1993). Stability of babyfaceness and attractiveness across the life span. Journal of Personality and Social Psychology, 64, 453-466.

Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-eyed" and "crooked-faced": Determinants of perceived and real honesty across the life span. Personality and Social Psychology Bulletin, 22, 1258-1269.

Appendices

Appendix 1 - The Sexual Serial Killers

NAME: Kenneth Bianchi
SOURCE: Jessel, Wilson, Morton and Waddell (1990a)
NUMBER OF VICTIMS: 14 over a period of 13 months.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse, sexual assault, victim lack of attire.

NAME: Ian Brady
SOURCE: Hall (1993)
NUMBER OF VICTIMS: three over a period of three years.
EVIDENCE OF SEXUAL MOTIVATION: sadistic torture, victim lack of attire.

NAME: Theodore Bundy
SOURCE: Jackson (1995)
NUMBER OF VICTIMS: More than 30 over a number of years.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse, sexual assault, victim lack of attire.

NAME: Angelo Buono
SOURCE: Jessel, Wilson et al. (1990a)
NUMBER OF VICTIMS: 12 over a period of four months.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse, sexual assault, victim lack of attire.

NAME: Harvey Carignan
SOURCE: Jessel, Wilson, Morton and Waddell (1991d)
NUMBER OF VICTIMS: five over a period of 25 years, 11 more suspected.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse, sexual assault, insertion of foreign object into body cavity, lack of attire.

NAME: Jerome Brudos
SOURCE: Jessel, Wilson, Morton and Waddell (1991b)
NUMBER OF VICTIMS: four over a period of 15 months.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse, sadistic torture, removal of sexual body parts, change of victim attire.

NAME: Jeffrey Dahmer
SOURCE: Jackson (1995)
NUMBER OF VICTIMS: 17 during one year.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse, sexual positioning of body, victim lack of attire.

NAME: Albert De Salvo
SOURCE: Jessel, Blake, Morton and Waddell (1990a)
NUMBER OF VICTIMS: 13 over a period of 18 months.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse, sexual assault, victim lack of attire, sexual positioning of body, insertion of foreign object into body cavity, exposure of sexual parts of body.

NAME: Westley Dodd
SOURCE: Lane and Gregg (1996)
NUMBER OF VICTIMS: three within a period of one year.
EVIDENCE OF SEXUAL MOTIVATION: sexual assault, victim lack of attire.

NAME: Gerald Gallego
SOURCE: Jessel, Wilson, Morton and Waddell (1990b)
NUMBER OF VICTIMS: 10 over a period of two years.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse.

NAME: Ed Gein
SOURCE: Black, Lane and Pender (1994a)
NUMBER OF VICTIMS: two, separated by a period of three years.
EVIDENCE OF SEXUAL MOTIVATION: sexual mutilation, removal of sexual body parts, fashioning clothing and other objects from body parts, victim lack of attire.

NAME: Harvey Glatman
SOURCE: Pender, Innes and Lane (1993b)
NUMBER OF VICTIMS: three over a one-year period.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse, victim lack of attire, change of victim attire.

NAME: Charles Hatcher
SOURCE: Ganey (1989)
NUMBER OF VICTIMS: several over a period of years.
EVIDENCE OF SEXUAL MOTIVATION: victim lack of attire, sexual assault, sexual intercourse.

NAME: Neville Heath
SOURCE: Jessel, Blake, Morton and Waddell (1990b)
NUMBER OF VICTIMS: two, separated by a two-week period.
EVIDENCE OF SEXUAL MOTIVATION: victim lack of attire, sexual mutilation, sadistic torture, insertion of foreign object into body cavity.

NAME: Paul Knowles
SOURCE: Jessel et al. (1991d)
NUMBER OF VICTIMS: 18 over a period of four months, suspected of up to 35.
EVIDENCE OF SEXUAL MOTIVATION: victim lack of attire, sadistic torture, sexual intercourse.

NAME: Earle Nelson
SOURCE: Lane and Gregg (1996)
NUMBER OF VICTIMS: at least 22 over 18 months.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse.

NAME: Danny Rolling
SOURCE: Lane and Gregg (1996)
NUMBER OF VICTIMS: five over two days, probably three more a year earlier.
EVIDENCE OF SEXUAL MOTIVATION: sexual mutilation, removal of sexual body parts, sexual positioning of body, victim lack of attire, sexual intercourse.

NAME: Arthur Shawcross
SOURCE: Jessel, Wilson, Morton and Waddell (1991a)
NUMBER OF VICTIMS: 13 over a period of 27 years.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse, victim lack of attire, exposure of sexual parts of body, sexual mutilation, removal of sexual body parts.

NAME: Frederick West
SOURCE: Gekoski (1998)
NUMBER OF VICTIMS: at least 10 over a period of 16 years.
EVIDENCE OF SEXUAL MOTIVATION: sexual intercourse, sexual assault, sadistic torture, victim lack of attire.

Appendix 2 - Details Identifying those Photographs from which the Preliminary and Main Study Arrays were selected

Table A lists the 10 sexual serial killer photographs and 208 *prima facie* non-criminal photographs from which those of the preliminary and main study photographic arrays were selected. The sources from which the photographs were obtained are provided, along with the names and occupations of those portrayed, as identified in those sources.

Table A. *Details Identifying those Photographs from which the Preliminary and Main Study Arrays were selected*

Photograph	Name	Occupation	Source
ONE	Jeffrey Dahmer	Sexual serial killer	Jackson (1995)
TWO	Albert De Salvo	Sexual serial killer	Jessel, Blake et al. (1990a)
THREE	Charles Hatcher	Sexual serial killer	Ganey (1989)
FOUR	Neville Heath	Sexual serial killer	Jessel, Blake et al. (1990b)
FIVE	Frederick West	Sexual serial killer	Gekoski (1998)
SIX	Angelo Buono	Sexual serial killer	Jessel, Wilson et al. (1990a)
SEVEN	Theodore Bundy	Sexual serial killer	Jackson (1995)
EIGHT	Ed Gein	Sexual serial killer	Black, et al. (1994a)
NINE	Harvey Carignan	Sexual serial killer	Jessel et al. (1991d)
TEN	Ian Brady	Sexual serial killer	Hall (1993)
#11	David Jessel	Journalist	Jessel, Blake, Morton and Waddell (1989a)
#12	James Morton	Lawyer	Jessel et al. (1989a)
#13	Bill Waddell	Policeman	Jessel et al. (1989a)
#14	William Saxbe	Attorney General	Jessel et al. (1991d)
#15	Charles E. Campbell	Highway Patrolman	Jessel et al. (1991d)
#16	Ron Angel	GBI agent	Jessel et al. (1991d)
#17	Anonymous	Policeman	Jessel et al. (1991d)
#18	Anonymous	Policeman	Jessel et al. (1991d)
#19	Archie Sonenstahl	Policeman	Jessel et al. (1991d)
#20	Ron Smith	Businessman/father of victim	Black, Lane and Pender (1994b)
#21	Ronald Marc George	Judge	Jessel, Wilson et al. (1990a)
#22	Donald Skepper	Postmaster/victim	Pender, Innes and Lane (1993a)
#23	Derek Astin	Postmaster/victim	Pender et al. (1993a)
#24	Christmas Humphreys	Lawyer	Jessel, Blake, Morton and Waddell (1990d)
#25	Ludovic Kennedy	Unspecified	Jessel et al. (1990d)
#26	Anonymous	Policeman	Jessel, Blake, Morton and Waddell (1990c)
#27	Douglas Stewart	Surviving victim	Jessel, Blake et al. (1990c)
#28	Carl Stotter	Surviving victim	Jessel, Blake et al. (1990c)
#29	Paul Nobbs	Surviving victim	Jessel, Blake et al. (1990c)
#30	William Whitelaw	Home Secretary	Jessel, Blake, Morton and Waddell (1989b)
#31	Ronald Gregory	Policeman	Jessel et al. (1989b)
#32	Anonymous	Unspecified	Jessel et al. (1989b)
#33	James Deagan	Father of victim	Jessel, Wilson, Morton and Waddell (1991e)
#34	Abner Cunningham	Policeman	Jessel et al. (1991e)
#35	Unknown	Unknown	Jessel et al. (1991e)

Photograph	Name	Occupation	Source
#36	Robert McNeil	Policeman	Jessel, Wilson et al. (1990c)
#37	Peter Walker	Victim	Pender, Smyth and Meenan (1994a)
#38	Christopher Dunn	Librarian/victim	Pender et al. (1994a)
#39	Anonymous	Policeman	Pender et al. (1994a)
#40	Perry Bradley	Businessman/victim	Pender et al. (1994a)
#41	Ken John	Policeman	Pender et al. (1994a)
#42	Edmund Duff	Retired soldier/victim	Pender, Smyth and Meenan (1994b)
#43	L. Murphy	Policeman	Jessel, Wilson, Morton and Waddell (1991c)
#44	Anonymous	Policeman	Jessel et al. (1991c)
#45	Ronald Lee Reidenhour	Witness	Hall (1993)
#46	Hugh Thompson	Soldier	Hall (1993)
#47	Ragnar Hagelin	Father of victim	Hall (1993)
#48	Bert Stillings	Policeman	Hall (1993)
#49	Jim Williams	Policeman	Hall (1993)
#50	Stefan Kiszko	Wrongly imprisoned	Hall (1993)
#51	Fred Anderson	Brother of victim	Hall (1993)
#52	John Bennett	Policeman	West and Hill (1995)
#53	Anonymous	Unspecified	Jackson (1995)
#54	Roman Polanski	Director/husband of victim	Jackson (1995)
#55	Michael Insco	Lawyer	Ganey (1989)
#56	Lee Nation	Lawyer	Ganey (1989)
#57	Joseph Holtslag	FBI agent	Ganey (1989)
#58	Henry Hudson	Lawyer	Mones (1995)
#59	Ray Williams	Policeman	Mones (1995)
#60	John Coale	Policeman	Mones (1995)
#61	Stephen Mardigan	FBI agent	Mones (1995)
#62	Peter Topping	Policeman	Topping and Ritchie (1989)
#63	Peter Timms	Reverand/prison governor	Topping and Ritchie (1989)
#64	Michael Fischer	Solicitor	Topping and Ritchie (1989)
#65	John Douglas	FBI agent	Douglas and Olshaker (1996)
#66	Charlie Chaplin	Actor	Eisenstaedt (1980)
#67	Alfred Cortot	Musician	Eisenstaedt (1980)
#68	Nathan Milsten	Musician	Eisenstaedt (1980)
#69	Yehudi Menukin	Musician	Eisenstaedt (1980)
#70	Vladimir Horowitz	Musician	Eisenstaedt (1980)
#71	Sergei Rachmaninoff	Composer	Eisenstaedt (1980)
#72	Bruno Walter	Conductor	Eisenstaedt (1980)
#73	Cecil Beaton	Photographer	Eisenstaedt (1980)
#74	Michael Farmer	Unspecified	Eisenstaedt (1980)
#75	Max Reinhardt	Director	Eisenstaedt (1980)
#76	Charles A. Lindbergh	Unspecified	Eisenstaedt (1980)
#77	Sinclair Lewis	Writer/Nobel laureate	Eisenstaedt (1980)
#78	Ernie Pyle	Journalist	Eisenstaedt (1980)
#79	Benjamin Britten	Composer	Eisenstaedt (1980)
#80	Arthur Hays Sultzburger	Publisher	Eisenstaedt (1980)
#81	Henry Ford II	Businessman	Eisenstaedt (1980)
#82	Andrew Wyeth	Artist	Eisenstaedt (1980)
#83	Edward E. Murrow	Radio & television reporter	Eisenstaedt (1980)
#84	William Franklin Graham	Baptist preacher	Eisenstaedt (1980)
#85	Roger Tory Peterson	Bird watcher	Eisenstaedt (1980)
#86	Alec Guinness	Actor	Eisenstaedt (1980)
#87	Lex Barker	Actor	Eisenstaedt (1980)
#88	Joe Pasternak	Producer	Eisenstaedt (1980)
#89	Moss Hart	Playwright/director	Eisenstaedt (1980)
#90	Gian Carlo Menotti	Composer	Eisenstaedt (1980)
#91	Dmitri Mitropoulos	Conductor	Eisenstaedt (1980)
#92	Dag Hammarskjold	Secretary to UN	Eisenstaedt (1980)

Photograph	Name	Occupation	Source
#93	Lester Bowles Pearson	UN President of General assembly	Eisenstaedt (1980)
#94	Henry Cabot Lodge	UN delegate	Eisenstaedt (1980)
#95	Dean Rusk	US Secretary of State	Eisenstaedt (1980)
#96	Lyndon B. Johnson	President of US	Eisenstaedt (1980)
#97	Mikhail Baryshnikov	Ballet dancer	Eisenstaedt (1980)
#98	Anonymous	Unspecified	Fink (1984)
#99 .	Anonymous	Unspecified	Fink (1984)
#100	Anonymous	Unspecified	Fink (1984)
#101	Anonymous	Unspecified	Fink (1984)
#102	Anonymous	Unspecified	Fink (1984)
#103	Anonymous	Unspecified	Fink (1984)
#104	Anonymous	Unspecified	Fink (1984)
#105	Anonymous	Unspecified	Fink (1984)
#106	Willem Mengelberg	Conductor	Eisenstaedt (1976)
#107	Carl Milles	Unspecified	Eisenstaedt (1976)
#108	Henry R. Luce	Unspecified	Eisenstaedt (1976)
#109	Thomas Hitchcock	Polo player	Eisenstaedt (1976)
#110	Jack Warner	Businessman	Eisenstaedt (1976)
#111	John Steuart Curry	Artist	Eisenstaedt (1976)
#112	Roy E. Larson	President of Time, Inc.	Eisenstaedt (1976)
#113	Whittaker Chambers	Editor	Eisenstaedt (1976)
#114	Odgen Pleissner	Artist	Eisenstaedt (1976)
#115	Edward Laning	Artist	Eisenstaedt (1976)
#116	David Fredenthal	Artist	Eisenstaedt (1976)
#117	Paul Yacqoulet	Artist	Eisenstaedt (1976)
#118	John Kenneth Galbraith	Unspecified	Eisenstaedt (1976)
#119	Gene Kelly	Actor	Eisenstaedt (1976)
#120	Hugh Casson	Unspecified	Eisenstaedt (1976)
#121	Joyce Cary	Unspecified	Eisenstaedt (1976)
#122	Charles Laughton	Unspecified	Eisenstaedt (1976)
#123	Lionel Trilling	Author	Eisenstaedt (1976)
#124	Leonard Bernstein	Composer/conductor	Eisenstaedt (1976)
#125	King Constantine	King of Greece	Eisenstaedt (1976)
#126	Walter Lipman	Journalist	Eisenstaedt (1976)
#127	Jacques Barzun	Unspecified	Eisenstaedt (1976)
#128	Samuel Barber	Composer	Eisenstaedt (1976)
#129	Noel Coward	Playwright	Eisenstaedt (1976)
#130	Willy Brandt	Mayor of Berlin	Eisenstaedt (1976)
#131	Arthur Watson	US Ambassador	Eisenstaedt (1976)
#132	A. Eisenstaedt	Photographer	Eisenstaedt (1976)
#133	J. William Fulbright	US Senator	Eisenstaedt (1976)
#134	Pierre Boulez	Conductor	Eisenstaedt (1976)
#135	Spiro Agnew	US Vice-president	Eisenstaedt (1976)
#136	John M. Mitchell	US Attorney General	Eisenstaedt (1976)
#137	Richard G. Kleindienst	US Deputy Attorney General	Eisenstaedt (1976)
#138	Melvin Laird	US Secretary of Defense	Eisenstaedt (1976)
#139	Ronald Ziegler	President Nixon's press secretary	Eisenstaedt (1976)
#140	H. R. Haldeman	White House Chief of Staff	Eisenstaedt (1976)
#141	Edwin Land	Inventor	Eisenstaedt (1976)
#142	Christiaan Barnard	Surgeon	Eisenstaedt (1976)
#143	Andre Previn	Unspecified	Eisenstaedt (1976)
#144	Tom Gormon	Tennis player	Eisenstaedt (1976)
#145	Andrew Macpherson	Assistant photographer	Snowdon (1983)
#146	Sebastian Coe	Olympic gold medallist	Snowdon (1983)
#147	Ian Charleston	Actor	Snowdon (1983)
#148	Claus Moser	Chairman of Royal Opera House	Snowdon (1983)
#149	Michael Holroyd	Biographer/author	Snowdon (1983)

Photograph	Name	Occupation	Source
#150	Barry Flanagan	Sculptor	Snowdon (1983)
#151	Frank Auerbach	Artist	Snowdon (1983)
#152	Carl Toms	Theatre designer	Snowdon (1983)
#153	Harold Evans	Editor	Snowdon (1983)
#154	Alfred Kazin	Unspecified	Morath (1986)
#155	Phillip Roth	Unspecified	Morath (1986)
#156	Bernard Buffet	Unspecified	Morath (1986)
#157	Paul Tortellier	Musician	Morath (1986)
#158	Marcus Prachensky	Unspecified	Morath (1986)
#159	Andre Malraux	Unspecified	Morath (1986)
#160	Tom Keogh	Unspecified	Morath (1986)
#161	Henry Moore	Sculptor	Morath (1986)
#162	Arnol Genthe	Photographer	Sobieszek and Irmas (1994)
#163	John F. Collins	Photographer	Sobieszek and Irmas (1994)
#164	Paul Outerbridge, Jnr.	Photographer	Sobieszek and Irmas (1994)
#165	Anonymous	Unspecified	Sobieszek and Irmas (1994)
#166	Floris Michael Neusiiss	Photographer	Sobieszek and Irmas (1994)
#167	Max Yavno	Photographer	Sobieszek and Irmas (1994)
#168	Robert Mapplethorpe	Photographer	Sobieszek and Irmas (1994)
#169	Timothy	Unspecified	Michals (1988)
#170	Duane Michals	Photographer	Michals (1988)
#171	Anonymous	Wrestler in a gym in Kiev	Michals (1988)
#172	Anonymous	Animal trainer	Michals (1988)
#173	Anonymous	Member of troupe	Michals (1988)
#174	Anonymous	Member of troupe	Michals (1988)
#175	Andy Warhol	Artist	Michals (1988)
#176	Christopher Isherwood	Unspecified	Michals (1988)
#177	Maxwell Caulfield	Unspecified	Michals (1988)
#178	Billy Woods	Unspecified	Michals (1988)
#179	Jules Styne	Unspecified	Michals (1988)
#180	Dan Mayers	Unspecified	Michals (1988)
#181	Michael Swain	Assistant photographer	Michals (1988)
#182	Keith Miller	Assistant photographer	Michals (1988)
#183	Paul Fortunato	Assistant photographer	Michals (1988)
#184	Justin Kimball	Assistant photographer	Michals (1988)
#185	Robert Lakow	Assistant photographer	Michals (1988)
#186	Jeff Mayfield	Assistant photographer	Michals (1988)
#187	Dave DeLuca	Assistant photographer	Michals (1988)
#188	Julian Jaimie	Assistant photographer	Michals (1988)
#189	Anonymous	Industrial engineer	Sander (1971)
#190	Anonymous	Student teacher	Sander (1971)
#191	Louis Ambrosi	Sculptor	Sander (1971)
#192	Otto Dix	Painter	Sander (1971)
#193	Jankel Adler	Painter	Sander(1971)
#194	Richard Seewald	Painter	Sander (1971)
#195	Anonymous	Artist	Sander (1971)
#196	Paul Rehkemper	Singer	Sander (1971)
#197	Paul Hindemith	Composer	Sander (1971)
#198	Leonardo Aramesco	Singer	Sander (1971)
#199	Gustav Hartung	Theatre director	Sander (1971)
#200	Albert Busche	Writer	Sander (1971)
#201	August Sander	Photographer	Sander (1971)
#202	Jeff	Unspecified	Ramsey (1999)
#203	Anonymous	Unspecified	Hart (1995)
#204	Anonymous	Unspecified	Premoli (1997)
#205	Anonymous	Unspecified	Premoli (1997)
#206	Peter O'Toole	Actor	Morley (1989)

Photograph	Name	Occupation	Source
#207	Anthony Powell	Unspecified	Morley (1989)
#208	Anonymous	Varnisher	Sander (1971)
#209	Sutzen	Lawyer	Sander (1971)
#210	Anonymous	Artist	Sander (1971)
#211	Anonymous	Unspecified	Sander (1971)
#212	Heine Hesse	Boxer	Sander (1971)
#213	Professor Abendroth	Academic	Sander (1971)
#214	Anonymous	Circus worker	Sander (1971)
#215	Billy J. Kramer	Unspecified	Morley (1989)
#216	Richard Ingrams	Journalist	Morley (1989)
#217	Tony Rushton	Journalist	Morley (1989)

Note: Photographs ONE to #66 were obtained from sources relating to serial murder. Photographs #67 to #189 were taken from anthologies of photographic portraits. Photographs #190 to #218 were selected from anthologies of photographic portraits and magazines to match certain SK photographs on age, grooming, lighting and picture quality.

Appendix 3 - Experimenter Ratings Used for Selecting the Photographs in the Preliminary Array

Table B displays the experimenter's ratings of the 218 photographs eligible for selection in the preliminary array for age, grooming, lighting, picture quality, the presence of environmental traces and facial orientation.

Table B. *Ratings of Age, Grooming, Lighting, Picture Quality, Environmental Traces and Orientation, made by the Experimenter for the Selection of Photographs in the Preliminary Array*

Photograph	Age	Grooming	Lighting	Picture Quality	Environmental Traces	Orientation
ONE	30s	2	4	4	absent	right
TWO	50s	3	3	4	absent	forward
THREE	40s	3	3	2	absent	forward
FOUR	30s	4	2	3	absent	forward
FIVE	40s	3	3	5	absent	forward
SIX	40s	3	3	4	absent	left
SEVEN	30s	3	3	1	absent	forward
EIGHT	50+	3	3	1	absent	forward
NINE	40s	3	3	3	absent	forward
TEN	20s	3	3	5	absent	forward
#11	40s	3	3	3	absent	forward
#12	40s	3	5	1	absent	forward
#13	40s	3	3	5	absent	forward
#14	50+	3	3	5	absent	forward
#15	20s	3	3	1	absent	forward
#16	40s	4	3	3	absent	left
#17	40s	3	3	3	absent	left
#18	50+	3	3	1	absent	right
#19	40s	3	5	1	absent	forward
#20	50+	3	3	5	absent	right
#21	50+	3	5	1	absent	right
#22	40s	5	5	1	absent	forward
#23	40s	3	3	1	absent	forward
#24	50+	5	3	3	absent	forward
#25	40s	3	3	3	absent	left
#26	50+	3	3	4	absent	forward
#27	40s	3	4	3	absent	forward
#28	20s	3	3	1	absent	forward
#29	20s	3	3	3	absent	right
#30	50+	3	3	1	absent	right
#31	50+	3	3	3	absent	right
#32	30s	3	3	1	absent	right
#33	50+	3	3	1	absent	right
#34	40s	3	3	4	absent	left
#35	50+	3	3	3	absent	left
#36	50+	3	3	5	absent	forward
#37	50+	3	5	1	absent	right
#38	40s	3	3	3	absent	forward

Photograph	Age	Grooming	Lighting	Picture Quality	Environmental Traces	Orientation
#39	30s	3	1	3	absent	forward
#40	30s	3	3	4	absent	forward
#41	30s	2	3	3	absent	forward
#42	40s	3	3	4	absent	forward
#43	50+	5	1	1	absent	forward
#44	40s	3	3	3	absent	left
#45	40s	3	1	1	absent	left
#46	30s	3	3	3	absent	forward
#47	20s	5	3	5	absent	forward
#48	40s	3	3	1	absent	forward
#49	30s	3	3	3	absent	forward
#50	30s	3	3	3	absent	right
#51	30s	3	1	1	absent	forward
#52	30s	3	3	3	absent	forward
#53	40s	3	1	1	absent	forward
#54	40s	5	3	3	absent	left
#55	40s	3	3	5	absent	left
#56	30s	3	3	3	absent	forward
#57	20s	3	3	2	absent	forward
#58	40s	3	3	5	absent	forward
#59	40s	3	3	3	absent	forward
#60	40s	3	3	5	absent	forward
#61	30s	3	3	1	absent	forward
#62	40s	3	3	3	absent	forward
#63	50+	3	3	5	absent	forward
#64	40s	3	3	5	absent	forward
#65	40s	3	3	5	absent	forward
#66	40s	3	3	3	absent	left
#67	40s	5	3	3	absent	left
#68	40s	5	3	5	absent	left
#69	40s	5	3	3	absent	left
#70	20s	3	3	3	absent	forward
#71	40s	5	3	3	absent	left
#72	50+	3	3	1	absent	forward
#73	40s	3	3	3	absent	left
#74	30s	5	3	3	absent	right
#75	30s	5	3	1	absent	right
#76	50+	3	3	1	absent	left
#77	40s	3	3	2	absent	forward
#78	50+	3	3	5	absent	right
#79	50+	3	3	5	absent	forward
#80	40s	3	3	3	absent	left
#81	50+	3	3	5	absent	forward
#82	30s	3	3	5	absent	forward
#83	40s	3	3	5	absent	forward
#84	50+	5	3	5	absent	forward
#85	30s	5	3	3	absent	forward
#86	50+	3	1	5	absent	forward
#87	30s	3	3	5	absent	forward
#88	20s	1	1	1	absent	forward
#89	50+	3	3	4	absent	forward
#90	40s	3	3	3	absent	forward
#91	40s	3	3	3	absent	forward
#92	50+	3	3	5	absent	forward
#93	40s	5	3	3	absent	left
#94	50+	3	3	3	absent	forward

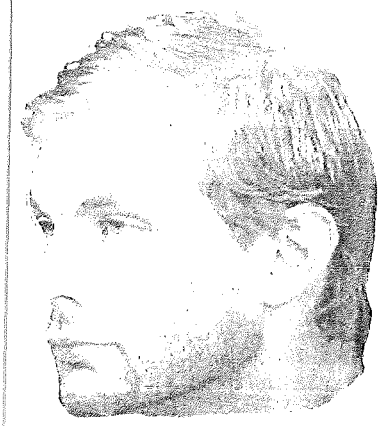
Photograph	Age	Grooming	Lighting	Picture Quality	Environmental Traces	Orientation
#95	50+	3	3	5	absent	forward
#96	50+	3	3	3	absent	right
#97	50+	3	3	4	absent	forward
#98	20s	3	3	5	absent	forward
#99	40s	3	3	5	absent	forward
#100	30s	3	3	5	absent	right
#101	20s	3	3	5	absent	left
#102	20s	3	3	5	absent	right
#103	20s	3	3	5	absent	forward
#104	20s	3	3	3	absent	forward
#105	20s	4	3	5	absent	forward
#106	50+	3	3	5	absent	forward
#107	50+	3	3	1	absent	forward
#108	50+	3	3	1	absent	right
#109	30s	3	5	1	absent	right
#110	30s	3	3	1	absent	left
#111	50+	3	1	5	absent	right
#112	40s	3	3	3	absent	left
#113	50+	3	3	5	absent	forward
#114	40s	3	3	4	absent	left
#115	40s	3	3	4	absent	forward
#116	30s	3	3	5	absent	forward
#117	30s	3	3	3	absent	forward
#118	50+	3	3	5	absent	forward
#119	40s	3	1	3	absent	forward
#120	30s	5	3	3	absent	forward
#121	50+	3	3	5	absent	forward
#122	50+	3	1	5	absent	right
#123	50+	1	3	5	absent	forward
#124	50+	3	3	5	absent	forward
#125	30s	1	2	5	absent	left
#126	20s	4	3	4	absent	forward
#127	50+	3	3	2	absent	forward
#128	40s	5	3	5	absent	left
#129	50+	3	3	5	absent	left
#130	50+	3	3	1	absent	left
#131	40s	3	3	2	absent	forward
#132	40s	5	3	3	absent	right
#133	50+	3	2	1	absent	forward
#134	50+	3	3	5	absent	forward
#135	40s	3	3	3	absent	forward
#136	50+	5	3	2	absent	forward
#137	50+	3	3	1	absent	forward
#138	50+	3	3	1	absent	forward
#139	50+	3	5	5	absent	forward
#140	30s	3	1	3	absent	forward
#141	40s	3	3	1	absent	right
#142	50+	3	3	3	absent	left
#143	40s	3	3	3	absent	left
#144	40s	3	3	5	absent	forward
#145	30s	3	3	3	absent	right
#146	20s	3	3	5	absent	forward
#147	20s	3	3	5	absent	forward
#148	30s	3	1	5	absent	forward
#149	40s	3	3	5	absent	forward
#150	40s	3	3	1	absent	left

Photograph	Age	Grooming	Lighting	Picture Quality	Environmental Traces	Orientation
#151	40s	3	1	3	absent	forward
#152	30s	3	3	5	absent	forward
#153	50+	3	3	5	absent	forward
#154	50+	3	3	5	absent	forward
#155	40s	3	3	5	absent	forward
#156	30s	3	3	4	absent	forward
#157	30s	5	3	3	absent	left
#158	30s	2	3	2	absent	right
#159	50+	1	3	5	absent	forward
#160	50+	3	1	3	absent	right
#161	50+	3	3	2	absent	forward
#162	50+	3	3	2	absent	forward
#163	50+	3	1	5	absent	forward
#164	40s	5	3	5	absent	left
#165	20s	4	3	5	absent	forward
#166	20s	3	3	4	absent	forward
#167	30s	3	3	5	absent	forward
#168	50+	3	3	2	absent	forward
#169	40s	3	3	4	absent	forward
#170	20s	3	3	3	absent	right
#171	30s	3	2	5	absent	forward
#172	30s	2	3	1	absent	forward
#173	50+	3	3	5	absent	forward
#174	20s	3	3	4	absent	right
#175	20s	3	3	4	absent	right
#176	20s	3	3	4	absent	forward
#177	50+	3	3	4	absent	forward
#178	20s	3	2	5	absent	left
#179	20s	3	2	4	absent	forward
#180	50+	3	3	3	absent	right
#181	20s	3	3	4	absent	forward
#182	20s	3	3	4	absent	forward
#183	20s	3	3	3	absent	forward
#184	20s	2	3	2	absent	left
#185	30s	3	3	4	absent	right
#186	20s	3	3	3	absent	forward
#187	30s	3	3	3	absent	forward
#188	20s	3	3	3	absent	forward
#189	20s	3	3	2	absent	left
#190	40s	3	3	3	absent	forward
#191	30s	3	3	3	absent	forward
#192	40s	2	3	3	absent	forward
#193	30s	4	3	4	absent	right
#194	40s	3	3	5	absent	forward
#195	40s	3	3	3	absent	forward
#196	40s	3	3	3	absent	forward
#197	30s	3	3	3	absent	forward
#198	30s	3	3	5	absent	forward
#199	40s	3	3	3	absent	forward
#200	40s	3	3	4	absent	forward
#201	40s	3	3	2	absent	forward
#202	50+	3	3	4	absent	forward
#203	30s	1	3	5	absent	right
#204	30s	3	3	4	absent	right
#205	30s	3	3	3	absent	right
#206	30s	2	3	4	absent	right

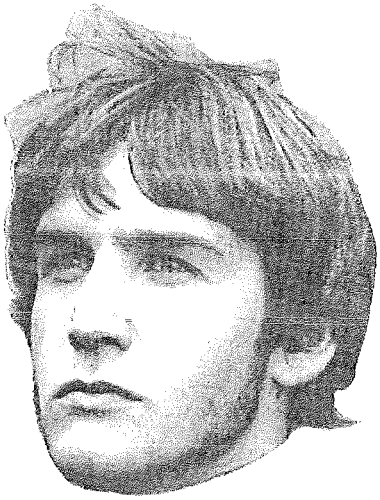
Photograph	Age	Grooming	Lighting	Picture Quality	Environmental Traces	Orientation
#207	30s	1	3	4	absent	right
#208	50+	3	3	3	absent	left
#209	30s	3	3	2	absent	forward
#210	30s	4	3	3	absent	forward
#211	30s	3	3	1	absent	forward
#212	40s	3	3	2	absent	forward
#213	30s	3	3	1	absent	forward
#214	30s	3	3	2	absent	forward
#215	30s	4	3	3	absent	forward
#216	30s	4	3	1	absent	forward
#217	30s	3	3	1	absent	forward
#218	30s	3	3	1	absent	forward

Appendix 4 - The Preliminary Photographic Array

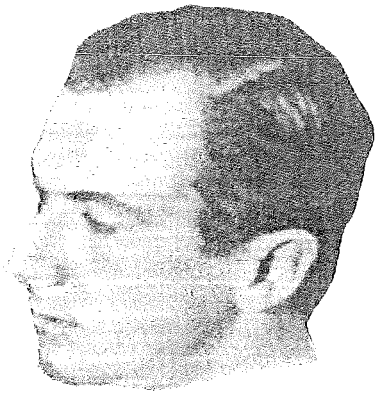
A copy of the preliminary photographic array follows below.



2



3

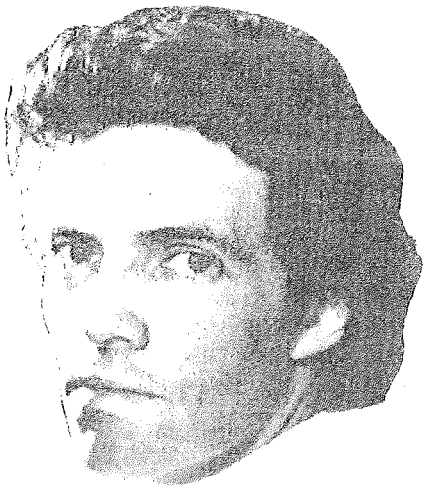


5

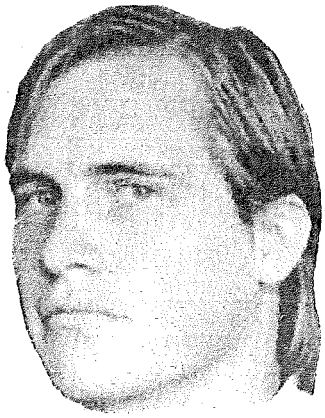
6

8

9



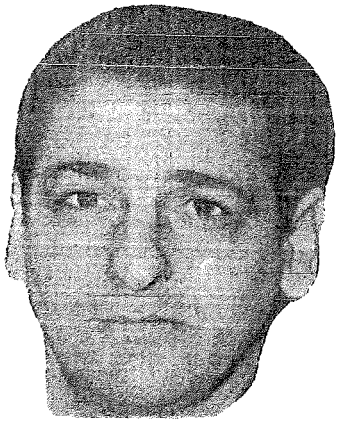
0



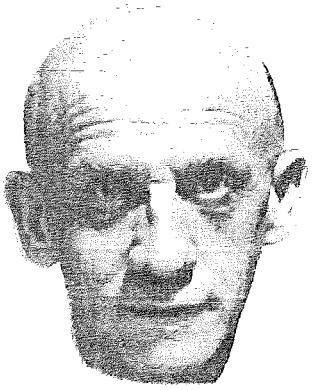
11

13

14



16

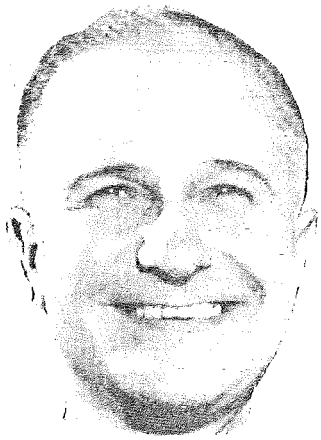


17

19

20

1



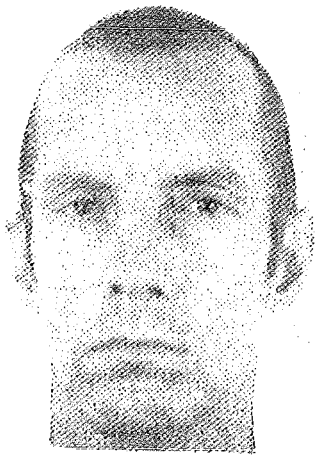
22

24

25

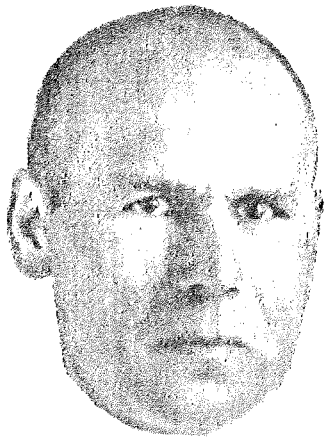
27

28

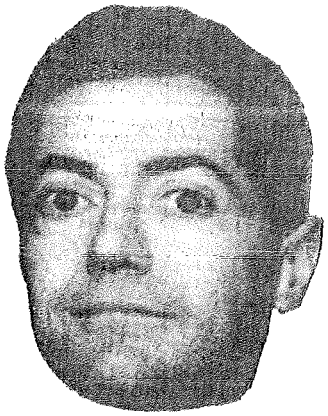


30

31



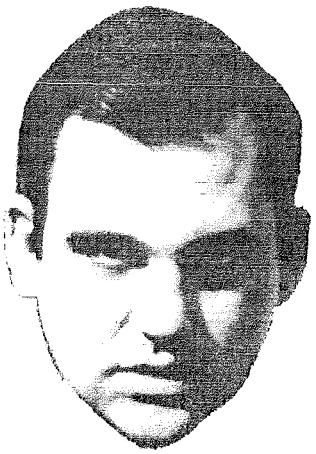
33



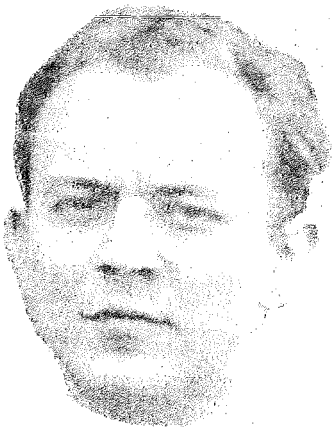
35

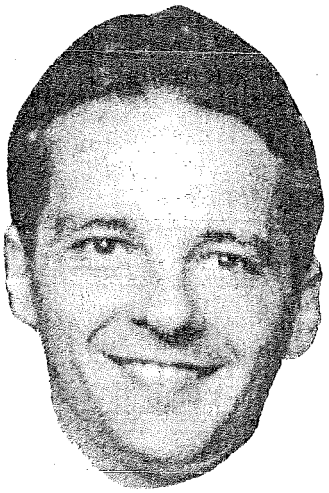
36

38



39

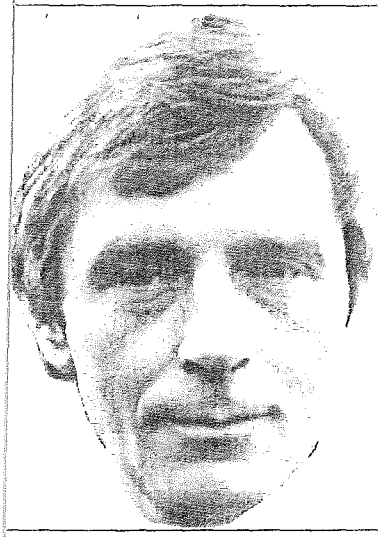


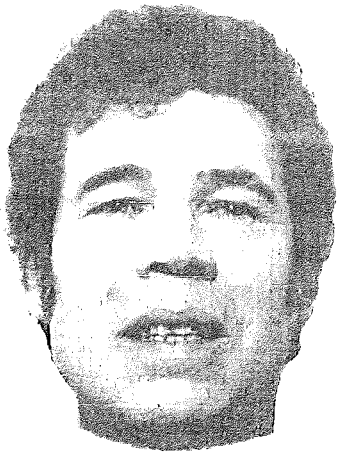


41

42

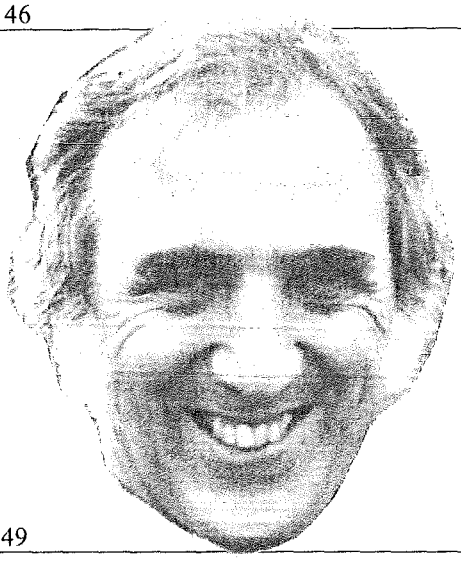
44



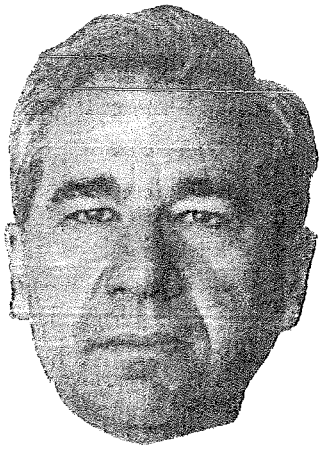


46

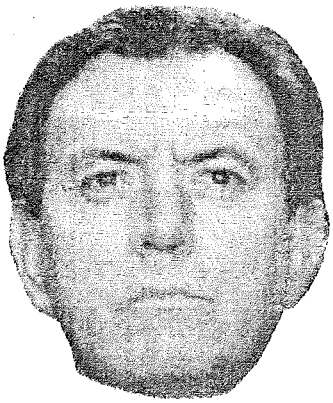
47



49



50

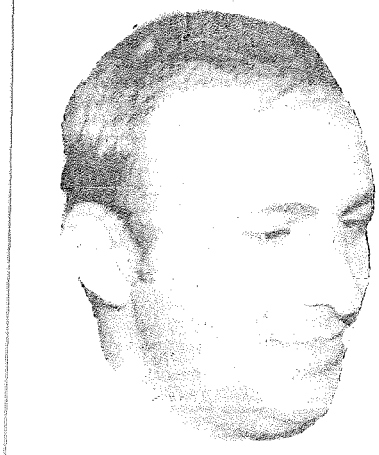


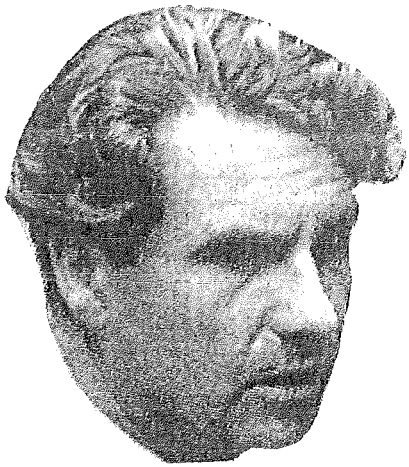
52

53

4

55



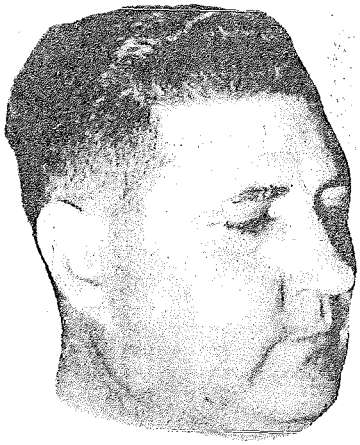


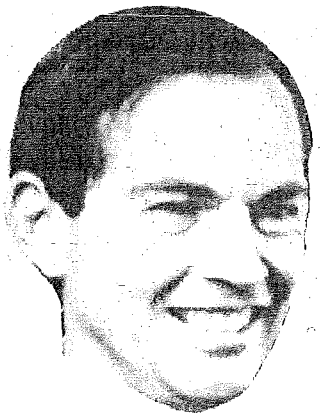
57

58

60

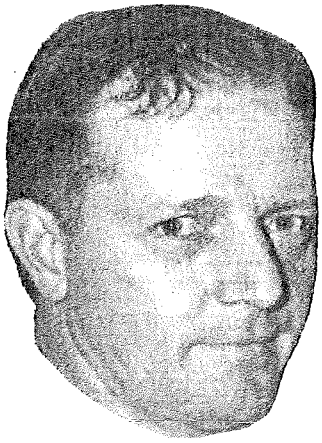
61



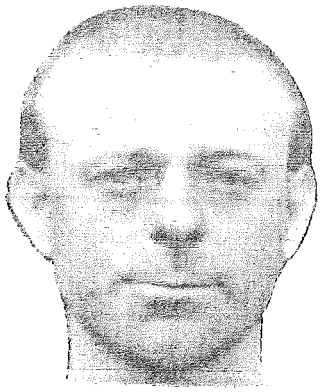


63

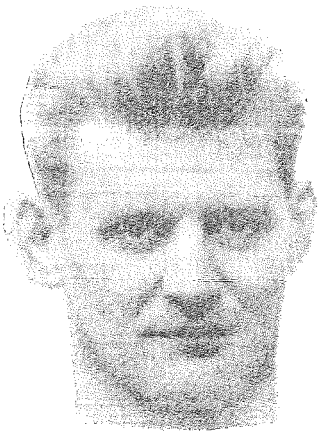
64



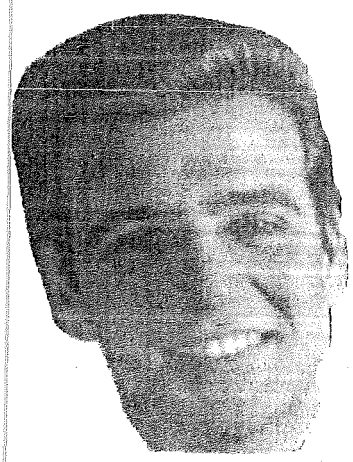
66



68

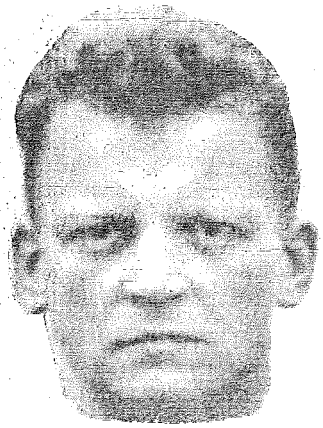


69



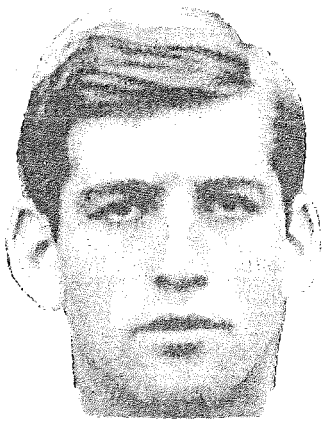
71

72



74

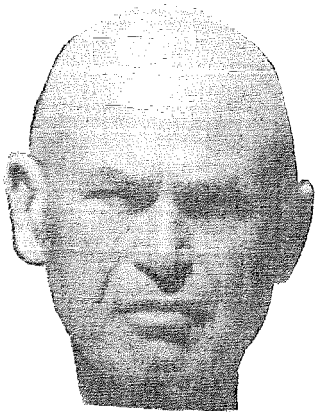
75



77

79

80

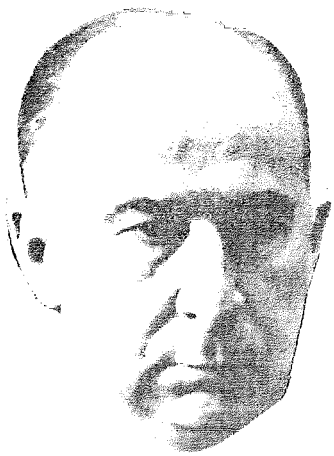


82

83

85

86



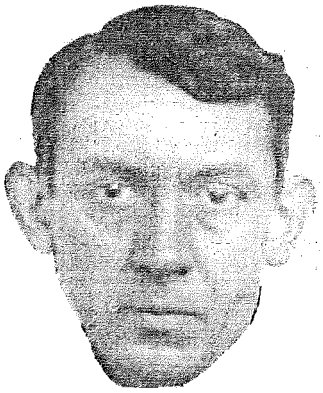
88

90

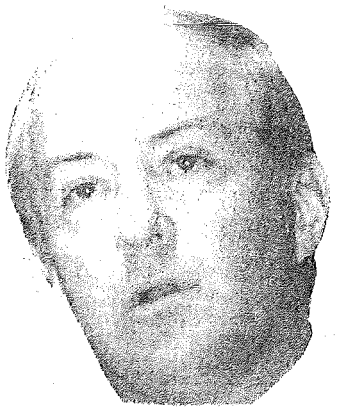
91

93

94

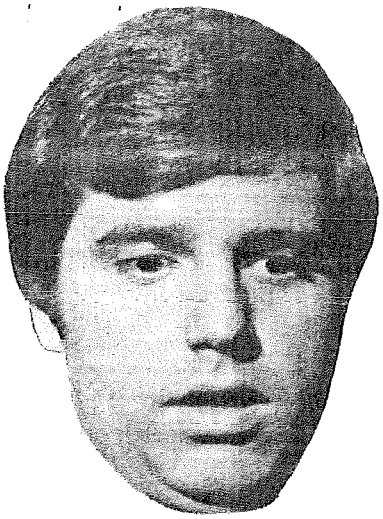


96

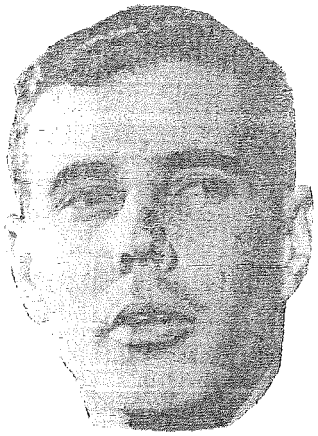


97

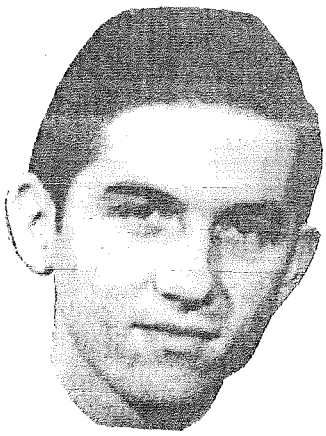
99



)

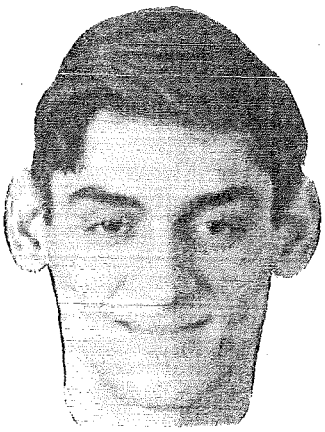


101

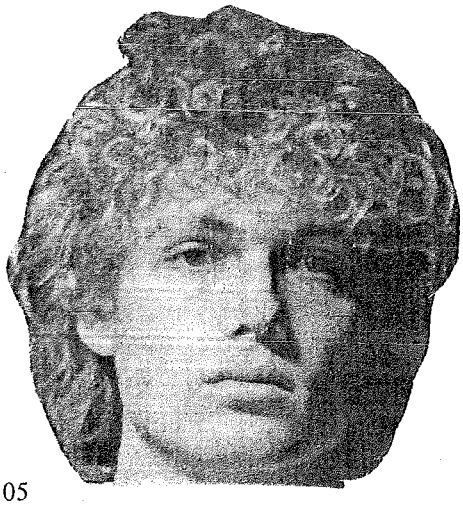


102

3



104

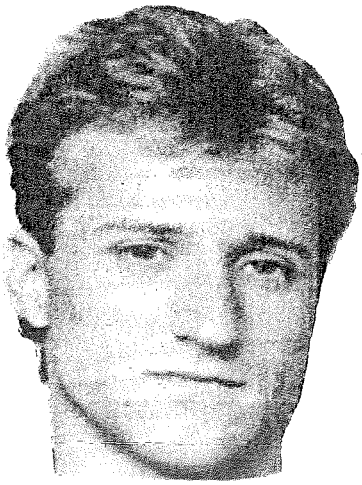


105

)6

107

108



09

110

Appendix 5 - The Preliminary Questionnaire

A copy of one version of the preliminary questionnaire follows below.

QUESTIONNAIRE

An investigation into accuracy in physiognomic perception: judgments of personality and behaviour from facial photographs

NOTE: You are invited to participate in the research project *An investigation into accuracy in physiognomic perception: judgments of personality and behaviour from facial photographs* by completing the following preliminary questionnaire. The aim of the project is *to investigate accuracy in person perception from facial photographs.* The aim of this preliminary questionnaire is to enable selection of photographs for the later experiment which are equivalent along certain confounding dimensions. The project is being carried out *as a requirement for an MSc* by *Richard Grundy*, who can be contacted at *342 9879,* under the supervision of *Drs. M. Barnes and L. Johnston.* He will be pleased to discuss any concerns you may have about participation in the project.

The questionnaire is anonymous, and you will not be identified as a participant without your consent. You may at any time withdraw your participation, including withdrawal of any information you have provided. **By completing the questionnaire, however, it will be understood that you have consented to participate in the project, and that you consent to publication of the results of the project with the understanding that anonymity will be preserved.**

TASK ONE

A. What is your sex? Please circle the correct response.

FEMALE MALE

B. The photographs labelled 1 to 110 on the accompanying sheets have been selected from published sources. You probably will **not** recognise any of them. Please look at these pictures, then list below the identifying numbers of any people who you are sure that you recognise and describe who you think they are.

TASK TWO

People can gather a lot of information from faces. Please rate **every** person (photographs 1 to 110) along **all** the following dimensions (A, B and C). Rate the photographs in the order given on this form and in the pages of photographs (note that the numbering is not necessarily consecutive).

An opposing description is placed at each end of a seven-point scale. Choose the score (1, 2, 3, 4, 5, 6, or 7) which best matches where you would place the person in reference to the opposing descriptions. For instance, a score of 4 would indicate that neither one nor the other opposing description better describes the person being rated. Do not spend too long on any one photograph as there are a lot of ratings for you to make. If you are unsure record your best guess. Write your chosen score (1 to 7) in the boxes provided for **every** photograph. Please do not leave gaps.

A. Facial expressions may be positive (e.g. happiness, excitement, pleasure, love, humour etc.) or negative (e.g. anger, sadness, disgust, fear, guilt, contempt etc.). This person's facial expression is:

very negative 1 2 3 4 5 6 7 very positive

1	2	3	4	5	6	7	8	9	10	11

12	13	14	15	16	17	18	19	20	21	22

23	24	25	26	27	28	29	30	31	32	33

34	35	36	37	38	39	40	41	42	43	44

45	46	47	48	49	50	51	52	53	54	55

56	57	58	59	60	61	62	63	64	65	66

67	68	69	70	71	72	73	74	75	76	77

78	79	80	81	82	83	84	85	86	87	88

89	90	91	92	93	94	95	96	97	98	99

100	101	102	103	104	105	106	107	108	109	110

B. This person is:

very physically unattractive 1 2 3 4 5 6 7 very physically attractive

1	2	3	4	5	6	7	8	9	10	11

12	13	14	15	16	17	18	19	20	21	22

23	24	25	26	27	28	29	30	31	32	33

34	35	36	37	38	39	40	41	42	43	44

45	46	47	48	49	50	51	52	53	54	55

56	57	58	59	60	61	62	63	64	65	66

67	68	69	70	71	72	73	74	75	76	77

78	79	80	81	82	83	84	85	86	87	88

89	90	91	92	93	94	95	96	97	98	99

100	101	102	103	104	105	106	107	108	109	110

C. This person has a:

very baby-like face 1 2 3 4 5 6 7 very mature face

1	2	3	4	5	6	7	8	9	10	11

12	13	14	15	16	17	18	19	20	21	22

23	24	25	26	27	28	29	30	31	32	33

34	35	36	37	38	39	40	41	42	43	44

45	46	47	48	49	50	51	52	53	54	55

56	57	58	59	60	61	62	63	64	65	66

67	68	69	70	71	72	73	74	75	76	77

78	79	80	81	82	83	84	85	86	87	88

89	90	91	92	93	94	95	96	97	98	99

100	101	102	103	104	105	106	107	108	109	110

Thank you for participating in this preliminary questionnaire.

Appendix 6 - Results of the Preliminary Study

The results from the preliminary study are reported below in Table C.

Table C. *Mean Ratings (N=10) of Physical Attractiveness, Babyfaceness and Expression for all the Photographs in the Preliminary Photographic Array, plus Within-Subject and Between-Subjects Difference Scores relating to Comparisons on these Variables between the Sexual Serial Killer Photographs and Matching Non-criminal Photographs*

Photograph	Matching SK Photograph	Mean Ratings of: PA	BAB	EXP	Within-Subject Difference Score	Between-Subject Difference Score
ONE	N/A	3.4	4.0	2.9	N/A	N/A
TWO	N/A	3.2	3.7	3.7	N/A	N/A
THREE	N/A	2.0	3.2	1.9	N/A	N/A
FOUR	N/A	2.0	2.8	2.3	N/A	N/A
FIVE	N/A	2.9	3.0	3.9	N/A	N/A
SIX	N/A	3.3	2.3	3.1	N/A	N/A
SEVEN	N/A	4.4	4.1	6.1	N/A	N/A
EIGHT	N/A	2.7	2.4	2.2	N/A	N/A
NINE	N/A	2.6	2.5	2.6	N/A	N/A
TEN	N/A	3.1	4.5	2.3	N/A	N/A
#11	NINE	4.4	4.3	4.5	1.83	1.97
#13	THREE	2.3	2.9	2.8	.90	.50
#16	SIX	2.0	2.5	2.5	.97	.70
#17	SIX	3.2	3.6	4.3	1.33	.87
#23	THREE	2.2	4.7	4.5	1.63	1.43
#25	SIX	2.9	2.7	3.0	.63	.30
#27	NINE	3.0	2.8	3.4	.77	.50
#34	SIX	3.1	3.1	3.7	.87	.53
#38	NINE	3.2	4.3	5.7	1.83	1.83
#39	FOUR	3.1	5.1	4.0	1.83	1.70
#42	FIVE	2.9	2.3	2.8	1.00	.60
#44	SIX	1.8	2.4	2.8	1.03	.63
#46	FOUR	2.8	4.3	2.9	1.10	.97
#47	TEN	4.6	4.5	3.7	1.50	.97
#48	THREE	2.5	3.4	3.6	1.00	.80
#49	FOUR	3.0	3.5	3.3	1.03	.90
#50	ONE	3.6	3.8	4.8	1.30	.77
#52	FOUR	3.5	4.4	2.6	1.30	1.13
#55	SIX	3.1	3.4	3.8	.93	.67
#58	FIVE	4.7	3.4	5.3	1.47	.87
#59	NINE	2.9	2.8	3.4	.93	.47
#60	THREE	2.7	3.5	4.5	1.60	1.20
#61	SEVEN	3.1	4.4	4.3	1.57	1.13
#62	NINE	4.2	4.1	5.5	2.03	2.03
#64	FIVE	3.6	2.6	5.9	1.37	1.03
#65	FIVE	3.7	2.5	4.1	1.07	.50
#70	TEN	3.7	5.0	3.8	1.00	.87
#72	EIGHT	1.6	1.6	2.5	.93	.73
#74	ONE	3.4	4.6	3.2	.83	.30

Photograph	Matching SK Photograph	Mean Ratings of: PA	BAB	EXP	Within-Subject Difference Score	Between-Subject Difference Score
#77	THREE	2.9	3.8	2.6	.87	.73
#80	SIX	2.6	2.2	3.2	.77	.30
#83	FIVE	3.3	2.4	4.3	1.20	.47
#84	TWO	2.3	1.5	2.6	1.53	1.40
#85	FOUR	3.1	3.1	3.6	1.03	.90
#90	NINE	2.4	3.2	5.7	1.80	1.33
#91	NINE	3.9	3.4	3.4	1.00	1.00
#92	TWO	1.6	2.0	3.3	1.30	1.23
#94	TWO	2.5	2.0	3.4	1.03	.90
#95	TWO	3.3	3.5	5.6	1.13	.73
#97	TWO	3.1	2.5	5.0	1.00	.87
#98	TEN	3.8	5.6	3.0	1.13	.83
#100	ONE	3.6	4.4	5.7	1.53	1.13
#105	TEN	3.8	4.5	3.2	1.13	.53
#107	EIGHT	1.6	2.2	2.2	1.17	.43
#112	SIX	1.6	2.1	1.9	1.23	1.03
#114	SIX	2.5	2.8	2.3	.83	.70
#115	NINE	2.4	2.9	2.9	.73	.27
#117	FOUR	2.7	4.4	2.6	1.13	.87
#118	TWO	3.5	3.2	5.9	1.27	1.00
#121	TWO	3.0	2.2	2.8	1.00	.87
#126	TEN	3.8	4.7	3.8	1.00	.80
#127	EIGHT	2.4	1.7	2.9	.97	.57
#131	THREE	2.6	2.4	3.3	1.07	.93
#133	EIGHT	2.4	2.1	4.4	1.40	.93
#135	NINE	3.3	3.4	4.6	1.20	1.20
#136	EIGHT	3.4	2.6	5.9	1.93	1.53
#137	EIGHT	2.6	1.8	2.7	.80	.40
#138	EIGHT	2.9	1.9	3.0	.70	.50
#140	FOUR	2.3	4.2	2.3	.83	.57
#143	SIX	3.4	4.6	5.5	1.93	1.60
#144	FIVE	3.2	3.7	3.8	1.17	.37
#145	ONE	4.1	3.9	3.5	.87	.47
#147	TEN	4.2	4.3	3.7	1.17	1.10
#149	FIVE	3.5	3.2	4.8	1.10	.57
#153	TWO	3.2	3.6	3.4	.87	.13
#155	FIVE	3.3	2.8	3.7	.87	.27
#156	FOUR	2.8	3.3	3.2	1.10	.73
#161	EIGHT	2.2	2.8	2.4	1.03	.37
#162	EIGHT	2.6	2.4	2.9	.73	.27
#166	TEN	2.6	4.6	4.9	1.53	2.17
#168	EIGHT	2.1	6.1	1.9	.80	.40
#169	FIVE	3.3	4.6	3.7	1.07	.73
#172	SEVEN	3.8	4.0	5.0	1.20	.60
#173	TWO	2.3	2.5	4.6	1.20	1.00
#176	TEN	3.6	5.0	3.0	.77	.57
#177	TWO	2.2	1.8	3.6	1.13	1.00
#181	TEN	4.8	5.6	3.3	1.47	1.27
#182	TEN	4.6	4.4	3.4	1.10	.90
#185	ONE	3.4	4.7	3.1	1.03	.30
#187	FOUR	4.0	4.2	5.7	2.33	2.27
#190	THREE	2.1	3.0	2.4	.53	.27
#194	THREE	2.4	4.3	2.7	.90	.77
#195	NINE	1.8	2.4	2.8	.77	.37
#196	THREE	1.9	3.1	3.6	1.17	.63
#197	FOUR	2.8	3.9	2.6	1.07	1.07

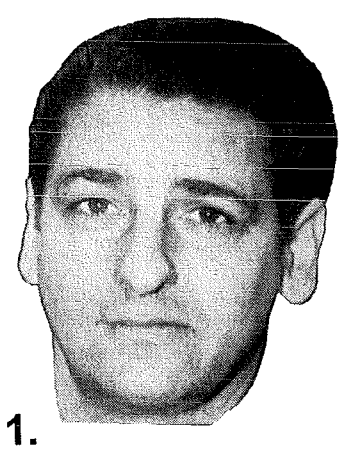
Photograph	Matching SK Photograph	Mean Ratings of: PA	BAB	EXP	Within-Subject Difference Score	Between-Subject Difference Score
#200	FIVE	3.4	3.5	3.8	1.10	.37
#201	THREE	2.8	4.1	3.3	1.37	1.03
#203	ONE	4.3	4.8	3.7	.97	.83
#204	ONE	4.5	4.5	3.2	1.03	.63
#205	ONE	6.0	5.0	5.8	2.23	2.17
#206	ONE	5.0	4.2	3.1	1.47	.67
#207	ONE	3.8	3.5	3.2	.73	.40
#209	SEVEN	2.8	4.0	3.8	1.60	1.33
#211	SEVEN	3.8	4.2	3.7	1.37	1.03
#213	SEVEN	2.5	4.2	2.8	2.10	1.77
#214	SEVEN	2.6	3.4	2.3	2.17	2.10
#215	SEVEN	4.1	4.3	3.6	1.33	1.00
#216	SEVEN	5.3	5.5	5.9	1.03	.83
#217	SEVEN	3.2	4.1	2.9	2.00	1.13
#218	SEVEN	2.8	3.6	2.7	2.03	1.83

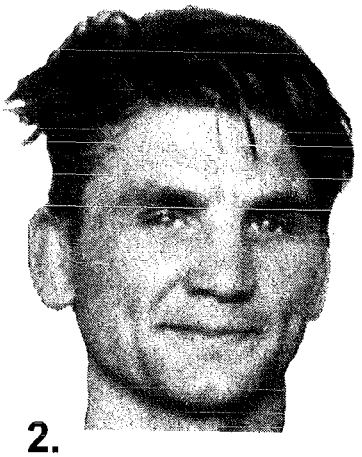
Photographs were matched according to ratings made by the experimenter relating to age, grooming, lighting, picture quality, environmental traces and orientation.
Babyfaceness is reverse scored so that a high value indicates a more baby-like face.
The within-subject difference score is the absolute difference between an individual participant's ratings of each non-criminal photograph and the matching SK photograph, averaged across PA, BAB and EXP. The between-subject difference score is the absolute difference between mean scores for each non-criminal photograph and the matching SK photograph, averaged across PA, BAB and EXP.

Appendix 7 - The Main Study Photographic Array

A copy of one version of the main study photographic array (Photographic Array Type "A") follows below.

PHOTOGRAPHIC ARRAY TYPE "A"


1.


2.

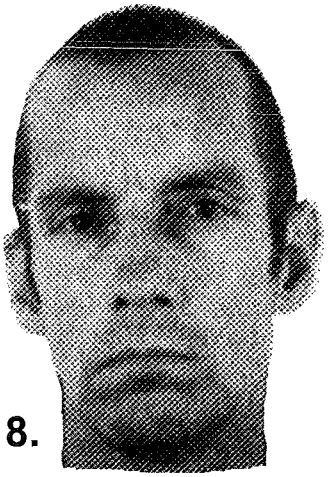

3.

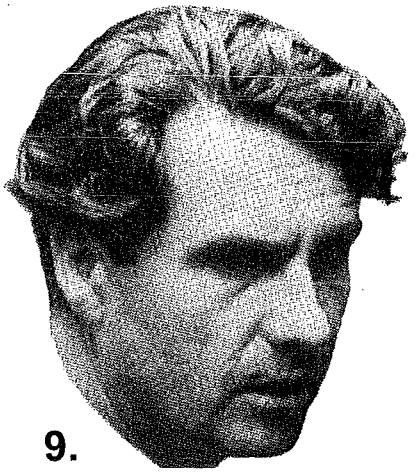

4.


5.

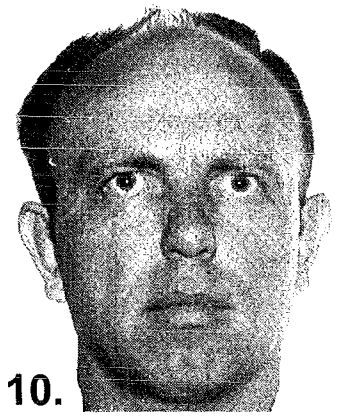

6.


7.

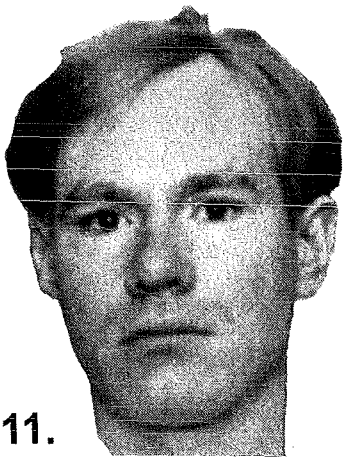

8.


9.

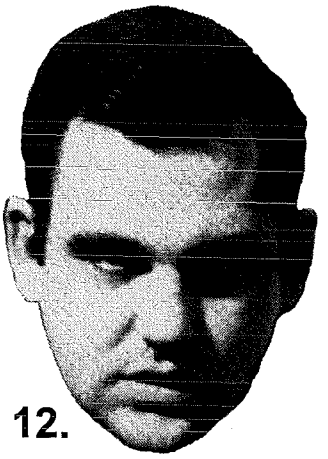
PHOTOGRAPHIC ARRAY TYPE "A"



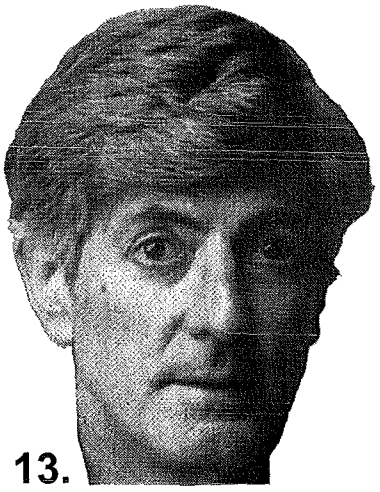
10.



11.



12.



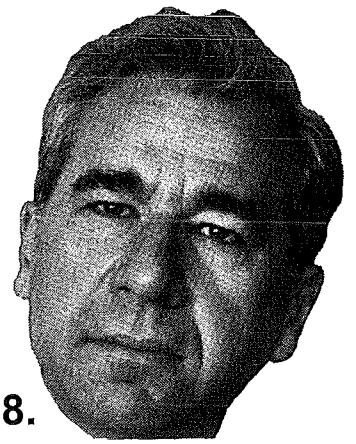
13.

14.

15.

16.

17.



18.

Appendix 8 - The Main Study Questionnaire

A copy of one version of the main study questionnaire (Type I) follows below.

QUESTIONNAIRE

An investigation into accuracy in physiognomic perception: judgments of personality and behaviour from facial photographs

NOTE: You are invited to participate in the research project *An investigation into accuracy in physiognomic perception: judgments of personality and behaviour from facial photographs* by completing the following questionnaire. The aim of the project is *to investigate accuracy in person perception from facial photographs.* The project is being carried out *as a requirement for an MSc* by *Richard Grundy*, who can be contacted at *342 3499,* under the supervision of *Drs. M. Barnes and L. Johnston.* He will be pleased to discuss any concerns you may have about participation in the project.

The questionnaire is anonymous, and you will not be identified as a participant without your consent. You may at any time withdraw your participation, including withdrawal of any information you have provided.
By completing the questionnaire, however, it will be understood that you have consented to participate in the project, and that you consent to publication of the results of the project with the understanding that anonymity will be preserved.

Do the following tasks in order. DO NOT LOOK AT TASK THREE UNTIL YOU HAVE COMPLETED ALL EARLIER TASKS. Do not refer back to tasks you have completed.

TASK ONE

A. What is your sex? Please circle the correct response. FEMALE MALE

What is your age? Please write your age in full years. AGE=_____

B. The title printed on the sheet of pictures I have received is (please fill in the gap):

PHOTOGRAPHIC ARRAY TYPE _____ (A, B, C, or D)

C. The photographs labelled 1 to 18 on the accompanying sheet have been selected from published sources. You probably will **not** recognise any of them. Please look at the photographs, then list below the identifying numbers of any people who you are sure that you recognise and describe who you think they are.

TASK TWO (TYPE I)

People can gather a lot of information from faces. Please rate **every** person (photographs 1 to 18) along **all** the following dimensions. An opposing description is placed at each end of a seven point scale. Choose the score (1, 2, 3, 4, 5, 6, or 7) which best matches where you would place the person in reference to the opposing descriptions. For instance, a score of 4 would indicate that neither one nor the other opposing description better describes the person being rated. Do not spend too long on any one photograph as there are a lot of ratings for you to make. If you are unsure record your best guess. Write your chosen score (1 to 7) in the boxes provided for **every** photograph. Please do not leave gaps.

1. This person's age is (write his age below in full years):

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

2. This person is:

very poorly groomed 1 2 3 4 5 6 7 very well groomed

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

3. This person:

has a very inflated opinion of himself 1 2 3 4 5 6 7 is very modest

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

4. This person is facing toward his:

extreme right 1 2 3 4 5 6 7 extreme left

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

5. This person has a:

very baby-like face 1 2 3 4 5 6 7 very mature face

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

6. This person:

does not feel guilt 1 2 3 4 5 6 7 feels guilt very easily

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

7. This person makes a habit of:

lying 1 2 3 4 5 6 7 telling the truth

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

8. This person is:

very manipulative 1 2 3 4 5 6 7 not at all manipulative

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

9. Facial expressions may be negative (e.g. anger, sadness, disgust, fear, guilt, contempt etc.) or positive (e.g. happiness, excitement, pleasure, love, humour etc.). This person's facial expression is:

very negative 1 2 3 4 5 6 7 very positive

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

10. This person is:

emotionally shallow 1 2 3 4 5 6 7 deeply emotional

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

11. The picture quality in this photograph is:

very poor 1 2 3 4 5 6 7 very good

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

12. This person's attitude to others is:

very uncaring 1 2 3 4 5 6 7 very caring

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

13. When this person is being charming he is:

very insincere and superficial 1 2 3 4 5 6 7 very genuine

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

14. This person accepts responsibility for his actions:

hardly ever 1 2 3 4 5 6 7 nearly always

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

15. This photograph is:

very dark 1 2 3 4 5 6 7 very light

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

16. This person is:

very physically unattractive 1 2 3 4 5 6 7 very physically attractive

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

Please make sure that you have completed all of Task Two fully. If you have finished, turn the page and complete Task Three. Do not refer back to Task Two after you have started reading Task Three.

DO NOT TURN THIS PAGE UNLESS YOU HAVE COMPLETED TASKS ONE AND TWO.

If you have not completed Tasks One and Two go back and finish those tasks.
DO NOT READ ANY FURTHER.

TASK THREE

Some of the photographs are actually of serious criminals and some are of non-criminals. Without referring back to your answers in Task Two, please rate each person on how likely they are to be a criminal.

very unlikely to be a criminal 1 2 3 4 5 6 7 very likely to be a criminal

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

Thank you for participating in this study.

Appendix 9 - Major Descriptive Statistics for All Variables and Photographs

The means, medians and standard deviations for the PCL-R derived variables, psychopathic traits, criminality and the physical variables for individual photographs are shown below in Table D.

Table D. *Major Descriptive Statistics for PCL-R Derived Variables, Psychopathic Traits, Criminality and Physical Variables for each Photograph (N=66)*

Variable	Photograph	Mean	Median	Std.Dev.
PCL-R derived variables				
Self-opinion (SEL)	TWO	3.98	4.00	1.32
	THREE	3.61	4.00	1.45
	FOUR	4.06	4.00	1.37
	FIVE	3.71	4.00	1.65
	SIX	3.35	3.00	1.58
	SEVEN	3.80	4.00	1.55
	EIGHT	3.86	4.00	1.38
	NINE	4.09	4.00	1.20
	TEN	3.42	3.00	1.55
	#25	3.61	4.00	1.41
	#115	3.58	3.00	1.40
	#140	3.26	3.00	1.74
	#153	4.91	5.00	1.39
	#155	3.80	3.50	1.65
	#162	4.59	5.00	1.55
	#172	3.12	3.00	1.51
	#176	3.91	4.00	1.36
	#190	4.05	4.00	1.33
Guiltiness (GUI)	TWO	4.01	4.00	1.48
	THREE	2.82	2.00	1.41
	FOUR	3.79	4.00	1.53
	FIVE	3.42	4.00	1.33
	SIX	3.73	4.00	1.21
	SEVEN	3.95	4.00	1.55
	EIGHT	3.42	3.00	1.63
	NINE	3.76	4.00	1.27
	TEN	3.29	3.00	1.55
	#25	4.20	4.00	1.47
	#115	4.15	4.00	1.23
	#140	2.23	2.00	1.52
	#153	5.08	5.00	1.29
	#155	4.09	4.00	1.45
	#162	4.62	5.00	1.42
	#172	3.27	3.00	1.39
	#176	4.11	4.00	1.67
	#190	3.86	4.00	1.35
Truthfulness (TRU)	TWO	3.92	4.00	1.59
	THREE	3.03	3.00	1.48

Variable	Photograph	Mean	Median	Std.Dev.
	FOUR	3.44	3.00	1.38
	FIVE	3.76	4.00	1.44
	SIX	3.76	4.00	1.34
	SEVEN	4.39	5.00	1.79
	EIGHT	3.15	3.00	1.33
	NINE	4.08	4.00	1.40
	TEN	3.03	3.00	1.47
	#25	3.89	4.00	1.43
	#115	4.11	4.00	1.38
	#140	2.17	2.00	1.12
	#153	5.53	6.00	1.36
	#155	4.48	5.00	1.58
	#162	4.65	5.00	1.42
	#172	3.82	4.00	1.53
	#176	3.98	4.00	1.47
	#190	4.30	5.00	1.69
Manipulative	TWO	4.05	4.00	1.69
tendencies (MAN)	THREE	3.30	3.00	1.46
	FOUR	3.79	4.00	1.32
	FIVE	3.82	4.00	1.48
	SIX	3.79	4.00	1.36
	SEVEN	4.06	4.00	1.64
	EIGHT	3.03	3.00	1.25
	NINE	4.06	4.00	1.39
	TEN	3.61	3.00	1.54
	#25	3.91	4.00	1.42
	#115	3.88	4.00	1.36
	#140	2.23	2.00	1.42
	#153	5.30	6.00	1.45
	#155	4.00	4.00	1.63
	#162	4.14	4.00	1.51
	#172	3.82	4.00	1.67
	#176	4.09	4.00	1.48
	#190	3.89	4.00	1.78
Emotional depth (EMO)	TWO	4.29	4.00	1.40
	THREE	3.24	3.00	1.50
	FOUR	3.42	3.00	1.67
	FIVE	4.14	4.00	1.49
	SIX	4.00	4.00	1.45
	SEVEN	4.47	5.00	1.43
	EIGHT	3.48	4.00	1.41
	NINE	3.52	4.00	1.15
	TEN	3.05	3.00	1.61
	#25	4.20	4.00	1.29
	#115	4.11	4.00	1.22
	#140	3.39	3.00	2.01
	#153	4.79	5.00	1.30
	#155	4.50	4.50	1.38
	162	4.55	5.00	1.42
	172	3.65	4.00	1.57
	176	4.06	4.00	1.43
	190	3.97	4.00	1.58
Attitude of care (CAR)	TWO	4.29	4.00	1.19
	THREE	2.89	3.00	1.35
	FOUR	3.18	3.00	1.19
	FIVE	3.82	4.00	1.32
	SIX	4.06	4.00	1.30

Variable	Photograph	Mean	Median	Std.Dev.
	SEVEN	4.95	5.00	1.40
	EIGHT	3.11	3.00	1.10
	NINE	3.74	4.00	1.19
	TEN	2.73	2.00	1.34
	#25	3.86	4.00	1.21
	#115	4.32	4.00	1.20
	#140	1.89	2.00	1.10
	#153	5.32	5.00	1.11
	#155	4.76	5.00	1.31
	#162	4.18	4.00	1.25
	#172	4.15	4.00	1.54
	#176	3.86	4.00	1.26
	#190	4.00	4.00	1.34
Sincerity of charm (CHA)	TWO	4.42	4.50	1.47
	THREE	3.36	3.00	1.42
	FOUR	3.91	4.00	1.31
	FIVE	3.68	4.00	1.59
	SIX	4.12	4.00	1.52
	SEVEN	4.76	5.00	1.73
	EIGHT	3.38	3.00	1.55
	NINE	4.29	4.00	1.31
	TEN	3.08	3.00	1.43
	#25	4.09	4.00	1.51
	#115	4.21	4.00	1.61
	#140	2.33	2.00	1.17
	#153	5.50	6.00	1.24
	#155	4.83	5.00	1.64
	#162	4.53	4.00	1.33
	#172	3.82	4.00	1.66
	#176	4.12	4.00	1.32
	#190	4.47	4.00	1.35
Responsibility (RES)	TWO	4.17	4.00	1.51
	THREE	3.12	3.00	1.43
	FOUR	3.83	4.00	1.45
	FIVE	3.53	3.00	1.44
	SIX	4.15	4.00	1.46
	SEVEN	4.65	5.00	1.41
	EIGHT	3.50	3.00	1.42
	NINE	4.29	4.00	1.30
	TEN	2.86	3.00	1.48
	#25	3.92	4.00	1.32
	#115	4.47	5.00	1.38
	#140	2.52	2.00	1.50
	#153	5.26	6.00	1.37
	#155	4.79	5.00	1.61
	#162	4.53	5.00	1.47
	#172	3.83	4.00	1.40
	#176	4.24	4.00	1.31
	#190	4.86	5.00	1.49
Main dependent variables				
Psychopathic traits (PSY)	TWO	30.86	30.00	8.33
	THREE	38.62	39.50	7.97
	FOUR	34.58	34.00	7.30
	FIVE	34.12	33.00	8.04
	SIX	33.05	33.00	7.14
	SEVEN	28.95	27.50	9.00

Variable	Photograph	Mean	Median	Std.Dev.
	EIGHT	37.06	37.00	7.17
	NINE	32.18	31.00	6.52
	TEN	38.94	40.00	8.42
	#25	32.32	32.00	6.94
	#115	31.18	31.00	7.12
	#140	43.98	44.00	6.66
	#153	22.32	21.00	7.15
	#155	28.74	28.50	9.00
	#162	28.21	28.00	7.78
	#172	34.52	34.00	9.10
	#176	31.62	31.00	7.72
	#190	30.59	31.00	7.36
Criminality (CRI)	TWO	4.21	4.00	1.55
	THREE	5.32	5.50	1.64
	FOUR	4.79	5.00	1.33
	FIVE	4.67	5.00	1.57
	SIX	3.73	4.00	1.47
	SEVEN	2.85	2.00	1.56
	EIGHT	5.33	5.00	1.19
	NINE	4.64	5.00	1.53
	TEN	4.97	5.00	1.55
	#25	3.88	4.00	1.40
	#115	3.47	3.00	1.48
	#140	5.91	6.00	1.45
	#153	2.45	2.00	1.57
	#155	2.98	2.00	1.76
	#162	3.64	4.00	1.58
	#172	3.55	3.00	1.58
	#176	4.21	4.00	1.51
	#190	3.41	3.00	1.52
Physical variables				
Age (AGE)	TWO	39.82	40.00	7.75
	THREE	29.23	28.50	5.59
	FOUR	33.15	32.50	6.05
	FIVE	37.44	38.00	6.96
	SIX	41.77	42.00	6.48
	SEVEN	40.08	38.00	7.07
	EIGHT	42.91	41.50	7.07
	NINE	38.32	38.00	6.22
	TEN	24.64	25.00	4.66
	#25	41.03	40.00	7.31
	#115	39.33	40.00	6.58
	#140	27.92	28.00	4.27
	#153	46.95	47.50	6.18
	#155	50.41	50.00	5.53
	#162	52.52	53.00	5.44
	#172	27.83	26.00	5.96
	#176	26.58	26.00	4.96
	#190	42.47	41.50	7.08
Grooming (GRO)	TWO	5.08	5.00	1.32
	THREE	3.62	4.00	1.41
	FOUR	4.88	5.00	1.14
	FIVE	2.92	3.00	1.35
	SIX	4.08	4.00	1.61
	SEVEN	5.55	6.00	1.04
	EIGHT	3.50	3.00	1.26
	NINE	3.88	4.00	1.17

Variable	Photograph	Mean	Median	Std.Dev.
	TEN	4.38	4.00	1.31
	#25	4.88	5.00	1.36
	#115	5.56	6.00	1.05
	#140	5.14	5.00	1.20
	#153	5.02	5.00	1.32
	#155	5.79	6.00	.85
	#162	4.92	5.00	1.22
	#172	4.09	4.00	1.69
	#176	4.12	4.00	1.12
	#190	5.74	6.00	.90
Orientation (ORI)	TWO	3.88	4.00	1.00
	THREE	3.88	4.00	.64
	FOUR	3.98	4.00	.57
	FIVE	4.07	4.00	.51
	SIX	4.64	5.00	2.03
	SEVEN	4.38	5.00	1.57
	EIGHT	4.82	5.00	1.16
	NINE	4.02	4.00	.37
	TEN	4.06	4.00	.43
	#25	4.62	5.00	2.04
	#115	4.21	5.00	1.76
	#140	3.89	4.00	.56
	#153	4.29	4.00	.92
	#155	3.64	3.00	1.42
	#162	4.23	4.00	1.17
	#172	4.21	4.00	.97
	#176	3.73	4.00	.87
	#190	3.95	4.00	1.33
Babyfaceness (BAB)	TWO	3.65	4.00	1.32
	THREE	3.35	3.00	1.03
	FOUR	4.12	4.00	1.35
	FIVE	3.17	3.00	1.12
	SIX	2.24	2.00	.93
	SEVEN	3.11	3.00	1.04
	EIGHT	2.70	3.00	1.19
	NINE	3.23	3.00	1.12
	TEN	5.41	5.00	1.08
	#25	2.74	2.50	1.22
	#115	3.77	4.00	1.42
	#140	4.92	5.00	1.45
	#153	2.82	3.00	1.39
	#155	2.27	2.00	1.06
	#162	1.95	2.00	1.07
	#172	4.94	5.00	1.40
	#176	5.15	5.00	1.17
	#190	3.06	3.00	1.54
Expression (EXP)	TWO	3.92	4.00	1.09
	THREE	2.35	2.00	.73
	FOUR	2.64	3.00	.97
	FIVE	4.20	5.00	1.35
	SIX	3.77	4.00	1.00
	SEVEN	6.53	7.00	.77
	EIGHT	2.74	3.00	1.09
	NINE	3.59	4.00	.94
	TEN	2.48	2.50	.98
	#25	2.88	3.00	.89
	#115	4.12	4.00	.98

Variable	Photograph	Mean	Median	Std.Dev.
	#140	1.77	1.00	1.02
	#153	4.38	4.00	1.00
	#155	4.12	4.00	1.07
	#162	3.33	3.00	.92
	#172	5.44	5.00	1.05
	#176	3.38	3.00	1.03
	#190	2.89	3.00	.90
Picture Quality (QUA)	TWO	4.74	5.00	1.33
	THREE	3.06	3.00	1.01
	FOUR	4.41	4.00	1.48
	FIVE	4.29	4.00	1.21
	SIX	4.77	5.00	1.40
	SEVEN	3.58	3.00	1.18
	EIGHT	2.53	2.00	1.06
	NINE	4.00	4.00	1.19
	TEN	4.95	5.00	1.34
	#25	3.32	3.00	1.13
	#115	5.08	5.00	1.14
	#140	3.05	3.00	1.14
	#153	4.53	5.00	1.27
	#155	5.20	5.00	1.18
	#162	3.35	3.00	1.07
	#172	4.55	4.00	1.24
	#176	3.98	4.00	1.27
	#190	5.26	6.00	1.15
Lighting (LIG)	TWO	4.73	5.00	1.14
	THREE	3.26	3.00	1.06
	FOUR	3.52	3.50	1.06
	FIVE	3.26	3.00	1.19
	SIX	4.24	4.00	1.11
	SEVEN	4.12	4.00	1.12
	EIGHT	3.26	3.00	1.27
	NINE	3.56	4.00	.99
	TEN	4.36	4.00	.89
	#25	3.94	4.00	1.18
	#115	4.44	4.00	1.15
	#140	2.85	2.50	1.43
	#153	3.26	3.00	1.36
	#155	3.67	4.00	1.11
	#162	2.88	3.00	1.05
	#172	4.09	4.00	1.03
	#176	4.15	4.00	1.07
	#190	4.29	4.00	.99
Physical Attractiveness (PA)	TWO	3.47	3.00	1.37
	THREE	3.06	3.00	1.26
	FOUR	2.82	3.00	1.15
	FIVE	2.71	2.00	1.33
	SIX	2.95	3.00	1.23
	SEVEN	4.80	5.00	1.29
	EIGHT	2.88	3.00	1.13
	NINE	3.15	3.00	1.13
	TEN	2.94	3.00	1.29
	#25	3.65	4.00	1.25
	#115	3.53	3.00	1.21
	#140	2.91	3.00	1.40
	#153	3.56	4.00	1.34
	#155	4.14	4.00	1.30

Variable	Photograph	Mean	Median	Std.Dev.
	#162	3.38	3.00	1.03
	#172	4.89	5.00	1.30
	#176	3.71	4.00	1.31
	#190	3.83	4.00	1.24

Note: Babyfaceness and the aggregate score for the psychopathic traits are reverse scored so that a high value indicates a more baby-like face and a more psychopathic personality.