Spot the Violent Sex Offender: Identifying Violent and Less-Violent Sex Offenders at a Glance

by

Mari Caitlin Shanahan Somerville

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Psychology

Carleton University

Ottawa, Ontario

© 2017

Mari Caitlin Shanahan Somerville

Abstract

People are able to make accurate social judgements from minimal information. Although person perception is believed to be an evolutionarily-evolved ability, not everyone is equally accurate. For instance, individuals scoring higher on psychopathic traits are better able to identify vulnerability than those who score lower. This study investigated which personality traits, including psychopathy, were associated with increased accuracy in perceiving violence and what facial cues individuals used when making their judgments. Two separate samples of undergraduates ($n = 267$ and $n = 203$) judged the potential for violence of 84 sexual offenders, from just two-second glances at facial photographs, and completed self-report measures of general personality dimensions and psychopathic traits. Contrary to prediction, participants in both samples were unable to distinguish between the violent and less-violent groups of offenders. However, as predicted, there were no gender differences in accuracy and some cues to masculinity tended to be valid cues to violence. Across both samples, most general personality characteristics measured by the HEXACO did not contribute to accuracy. However, under the more controlled conditions for sample two, low honesty-humility, a characteristic which has previously been associated with increased psychopathy, contributed to greater accuracy. Moreover, boldness and total psychopathy, as defined by the Triarchic Psychopathy Measure (TriPM), both contributed to increased accuracy in the detection of violence for sample two. Taken together, the results suggest that people use both valid and misleading cues when making quick judgments about others' potential for violence and that the relationship between psychopathic personality traits and accuracy in threat detection warrants further investigation.

Acknowledgements

I would first like to thank Carleton University. In my years here, my mind has expanded and I have grown in ways that I did not anticipate. For this, and for affording me the opportunity to meet some of the most amazing people, I will be forever grateful. Next, I would like to extend a huge thank you to my supervisor, Dr. Adelle Forth, for her ongoing support throughout my Master's. She is an inspiring individual and I feel lucky to have had the opportunity to work with her. She has allowed me to work independently to make this thesis my own but has always been available and willing to share her knowledge and advice when it was needed – not to mention, to read countless (lengthy) drafts. Thank you, Adelle, for your encouragement and wisdom and, above all, for playing a major role in my development as a graduate student and a researcher.

To Dr. John Weekes and Dr. Joanna Pozzulo, thank you for taking the time to read and critique my initial study proposal to help me develop it into a final, cohesive product. Your thoughtful insights led me to think about my study from new perspectives and have undoubtedly strengthened my thesis in many ways. Thank you to Dr. Ralph Serin for stepping in for Dr. Pozzulo as a committee member for my final defence as she begins her sabbatical and to Dr. Steven Prus for agreeing to be my external examiner. I appreciate the time that all of you have taken out of your busy schedules to take part in the final stage of my Master's career. I would also like to extend my gratitude to Dr. Kevin Nunes, who was my honour's thesis supervisor, for agreeing to act as the chair for my final defence. Thank you, Kevin, for the role you have played in my post-secondary career and for sharing your time to be a part of the final stage of another one of my degrees.

Thank you to the amazing staff in the graduate psychology department. In particular, thank you to Etelle Bourassa, Lea Hamilton, and Jessica Whitney for all that you do to keep the

department running. I greatly appreciate your consistent willingness to help and to answer any and all questions that arise, as I am very sure that everyone in the department does.

Thank you to all of my volunteers for your willingness to help me with this project. Without you and the time you contributed, there would be no study. A special thank you to Mary Ritchie for walking the path before me and answering my many questions, sharing tips, and revising one of my very first section drafts. I will not forget your readiness to share your experience and wisdom. I would also like to thank Cecilia Jorgenson for being my statistics guru. Thank you for all your help, Cecilia.

Next, I would like to thank my friends for being a major part of my support system throughout this journey. In particular, thank you to Chloe Pedneault, my dear friend and colleague. I thank you sincerely for all the help, advice, support, and encouragement you have extended to me throughout the four years that I have known you. I am grateful to have made such a remarkable lifelong friend. Thank you also to Kyle Burrows, for the time you have taken to read lengthy drafts of this thesis and to offer thoughtful opinions and suggestions to help make it better. More importantly, thank you for being there in moments of uncertainty and offering support and encouragement when it was needed. Thank you, also, to my mother's husband, Dr. Ian Cameron, for the academic advice you have given me throughout my degree. It has been very helpful to me at various stages and I truly value the knowledge you have shared with me.

Finally, words cannot express how grateful I am to my inspiration, my hero, my greatest supporter, and my friend – my mother, Dr. Tina Shanahan. Without you, I know that I would not have made it this far. You are undoubtedly the strongest and wisest person I know and I am so grateful to you for the innumerable ways in which you have supported and encouraged me throughout the years. This is for you, Mum.

Table of Contents

Abstract	ii
Acknowledgments	iii
List of Tables	ix
List of Figures	xi
List of Appendices	xii
Introduction	1
The Evolution of Interpersonal Judgement	4
Gender Differences in Perception	4
Evolutionary Frameworks for Variations in Threat Detection	8
Personality	12
The Psychopathic Personality	12
The HEXACO Model of Personality	15
Research on Immediate Interpersonal Judgement	17
Perception of Dominance.....	20
Perception of Aggression and Violence.....	24
Perception of Psychopathy	27
Cues to Aggression and Violence	28
Testosterone and Facial Width-to-Height Ratio	29
Valid, Poor, and Misleading Facial Cues to Violence	32
The Ability of Psychopathic Individuals to Detect Non-Verbal Cues	34
Current Study	37
Method	45

Participants and Sampling	45
Measures	45
Sex Offender Photographs	45
Ratings of Non-Verbal Cues	46
Facial Width-to-Height Ratio	47
Demographic Questionnaire	47
Attention Filters	47
HEXACO-60	47
Self-Report Psychopathy Scale (SRP-4)	48
Brief Triarchic Psychopathy Measure (Brief TriPM)	49
Accuracy Measure	50
Procedure	50
Results	53
Primary and Secondary Datasets	53
Inter-Rater Reliability	53
Secondary Dataset	53
FWHR	54
Data Preparation: Sample One	55
Missing Data	55
Linearity and Homoscedasticity	56
Univariate Outliers and Normality	56
Multicollinearity	57
Descriptive Statistics and Internal Consistency	58

Primary Analyses: Sample One	61
Distinguishing Target Violence	61
Gender Differences in Accuracy	61
Valid, Poor, and Misleading Facial Cues	62
Valid Cues and Decision Making	67
Psychopathy and Accuracy	68
General Personality and Accuracy	69
Data Preparation: Sample Two	71
Missing Data	71
Linearity and Homoscedasticity	72
Univariate Outliers and Normality	72
Multicollinearity	73
Descriptive Statistics and Internal Consistency	73
Primary Analyses: Sample Two	77
Distinguishing Target Violence	77
Gender Differences in Accuracy	77
Valid, Poor, and Misleading Facial Cues	78
Valid Cues and Decision Making	84
Psychopathy and Accuracy	85
General Personality and Accuracy	87
Discussion	88
Distinguishing Target Violence	88
Gender Differences in Accuracy	89

Valid, Poor, and Misleading Facial Cues	91
Valid Cues	91
Poor Cues	96
Misleading Cues	97
Valid Cues and Decision Making	99
Psychopathy and Accuracy	100
General Personality and Accuracy	102
Implications	103
Limitations	106
Future Directions	110
Conclusion	112
References	113

List of Tables

Table 1. <i>HEXACO Dimensions and Corresponding HEXACO-PI Facets</i>	16
Table 2. <i>Valid, Poor, and Misleading Facial Cues Used in Judgements of Violence and Threat</i>	34
Table 3. <i>Internal Consistency Statistics for Inter-Rater Reliability Among Facial Raters</i>	54
Table 4. <i>Sample One: Descriptive Statistics and Internal Consistencies for the HEXACO-60, SRP-4, and TriPM</i>	58
Table 5. <i>Sample One: Bivariate Correlations of the SRP-4 and the TriPM with the HEXACO-60</i>	59
Table 6. <i>Sample One: Bivariate Correlations between the SRP-4 with the TriPM</i>	60
Table 7. <i>Sample One: Bivariate Correlations of Facial Cues with PLV and Actual Violent History</i>	64
Table 8. <i>Sample One: Valid, Poor, and Misleading Facial Cues</i>	66
Table 9. <i>Sample One: Bivariate Correlations between Each SRP-4 and TriPM Personality Dimension and Accuracy</i>	69
Table 10. <i>Sample One: Bivariate Correlations between Each HEXACO-60 Personality Dimension and Accuracy</i>	70
Table 11. <i>Sample Two: Descriptive Statistics and Internal Consistencies for the HEXACO-60, SRP-4, and TriPM</i>	74
Table 12. <i>Sample Two: Bivariate Correlations of the SRP-4 and the TriPM with the HEXACO- 60</i>	75
Table 13. <i>Sample Two: Bivariate Correlations between the SRP-4 and the TriPM</i>	76
Table 14. <i>Sample Two: Bivariate Correlations of Facial Cues with PLV and Actual Violent</i>	

<i>History</i>	80
Table 15. <i>Sample Two: Valid, Poor, and Misleading Facial Cues</i>	83
Table 16. <i>Valid, Poor, and Misleading Facial Cues Replicated across Samples One and Two</i>	83
Table 17. <i>Sample Two: Bivariate Correlations between Each SRP-4 and TriPM Personality</i> <i>Dimension and Accuracy</i>	86
Table 18. <i>Sample Two: Bivariate Correlations between Each HEXACO-60 Personality</i> <i>Dimension and Accuracy</i>	87

List of Figures

<i>Figure 1.</i> Counter-Adaptation Evolutionary ‘Arms Race’ Diagram	11
<i>Figure 2.</i> Four Facet Conceptualization of Psychopathy	13

List of Appendices

Appendix A <i>SONA Online Recruitment Bulletin</i>	134
Appendix B <i>Sex Offender Photographs</i>	135
Appendix C <i>Rating Scales of Non-Verbal Cues</i>	137
Appendix D <i>Demographic Questionnaire</i>	141
Appendix E <i>Attention Filters</i>	142
Appendix F <i>HEXACO-60</i>	143
Appendix G <i>Brief TriPM</i>	146
Appendix H <i>Informed Consent Form</i>	149
Appendix I <i>Debriefing Form</i>	151
Appendix J <i>Propensity for Violence Likert Scale</i>	153
Appendix K <i>Correlations of Cues with PLV by Gender and Actual Target Violence</i>	154

Spot the violent sex offender: Identifying violent and less-violent sex offenders at a glance

Without expending much cognitive energy (e.g., Carlston & Skowronski, 2005; Schaller, 2008), humans use non-verbal cues to make automatic social judgements about the personalities and characteristics of people they have just met (e.g., Martelli, Majib, & Pelli, 2005). These cues include individual characteristics and actions, such as facial expression, structure, or masculinity (e.g., Oosterhof & Todorov, 2008; Sell, Cosmides, Tooby, Sznycer, von Rueden, & Gurven, 2009; Stirrat & Perrett, 2015; Willis & Todorov, 2006), physical and interpersonal attractiveness (e.g., Fowler, Lilienfeld & Patrick, 2009), and inadvertent non-verbal behaviours (i.e., ‘leakages’ of true intentions or feelings; e.g., Furley & Schweizer, 2016; Pretsch, Flunger, Heckmann, & Schmitt, 2013; Richards, Rollerson, & Phillips, 1991). It has been posited that the ability to detect the characteristics and emotions of others is an adaptation that has evolved in human beings in order to enhance survival and reproductive success (e.g., McArthur & Baron, 1983). That is, from an evolutionary perspective, the adaptation of violence and hostility in humans made it necessary for counter-adaptations to evolve, such as the ability to detect threatening cues in others (Duntley & Schackelford, 2008).

Research suggests, however, that not everyone is equally accurate in their evaluations of other people. Some personality characteristics have been identified as providing an advantage to individuals in the domain of person perception (i.e., perceptions about the characteristics of other people). In light of these findings, researchers have conducted studies on victim vulnerability and victim selection with the goals of identifying how vulnerable individuals may non-verbally convey signals to their vulnerability (e.g., Grayson & Stein, 1981) and which assessor personality traits may be related to accuracy in perceiving vulnerability (e.g., Book, Costello, & Camilleri, 2013; Ritchie, 2016). The psychopathic personality is one that has been found to be

associated with more success in the identification of vulnerable people by using only non-verbal cues (e.g., Book, Quinsey, & Langford, 2007; Ritchie, 2014, 2016; Wheeler, Book, & Costello, 2009). Research suggests that psychopathic individuals may be more attuned to non-verbal cues that convey vulnerability than non-psychopathic individuals (Book et al., 2013; Book et al., 2007; Camilleri, Kuhlmeier, & Chu, 2010; Wheeler et al., 2009).

In contrast, very few studies have examined whether there are certain observer personality traits that are associated with greater accuracy in detecting individuals who may pose a threat. Examining relationships of this sort may provide novel information about how individuals perceive threats in their environments. To date, research in this domain has focused on the cues conveyed by individuals that relate to (a) observer evaluations of an increased propensity toward aggression or violence and (b) an actual history of aggressive or violent behaviour, rather than the characteristics of observers that may increase their accuracy in such evaluations (e.g., Carré, McCormick, & Mondloch, 2009; Stillman, Maner, & Baumeister, 2010). Thus, a focus on the observer personality characteristics that are associated with a greater accuracy in assessing violence propensity is warranted. Psychopathy may be among such characteristics, as research evidence suggests an enhanced ability of psychopathic individuals to detect vulnerability in others based on non-verbal cues (e.g., Book et al., 2013; Ritchie, 2016). Perhaps this enhanced perception also extends to a greater ability to detect threatening individuals.

Despite the extant research on the use of non-verbal cues in rapid judgements, otherwise known as ‘first impressions’, about other people, research has not yet examined whether there are certain general personality traits (e.g., agreeableness, extraversion) that enhance individuals’ accuracy in rapid judgements of violent propensity. Moreover, the success of psychopathic

individuals in the detection of threat has yet to be explored. As such, the purpose of this study was to examine the relationship between personality traits, including psychopathic personality traits, and accuracy in the identification of violent individuals. Furthermore, the ability of laypeople to make accurate judgements regarding violence based on very limited exposure to photographs of faces (i.e., non-verbal cues) was explored. The automaticity with which all participants, males and females, are able to make these perceptions should speak to whether automatic interpersonal judgements evolved in humans as a universal mechanism for threat detection. Furthermore, an ability to detect a propensity for violence simply by looking briefly at a face may provide insight into the types of threat-relevant cues, such as features associated with higher levels of testosterone, that come into play during rapid interpersonal judgements.

The relevant evolutionary frameworks that may help to explain variations in the accuracy of humans' immediate evaluations of others are reviewed first. Next, the common methodological design implemented to examine humans' capacity to immediately perceive aspects about others based solely on non-verbal cues is discussed. Further, the existing literature in this domain regarding humans' ability to perceive threat- and violence-relevant cues is reviewed, followed by an overview of the valid and misleading cues that people use in their judgements. The literature specifically concerning judgements of violence based on facial evaluations, including cues that have proven to be valid or misleading, is then explored in more depth. This review is followed by more detailed discussions of both psychopathy and the general personality dimensions of the HEXACO model of personality (i.e., **H**onesty-Humility, **E**motionality, **eX**traversion, **A**greeableness, **C**onscientiousness, **O**penness; Lee & Ashton, 2004). Research questions and hypotheses are described following the literature review.

The Evolution of Interpersonal Judgement

A number of researchers have adopted an evolution-centred approach to explaining why humans may be able to quickly and accurately detect the characteristics of other people (e.g., Damasio, 1994; de Waal, 2007; Frith & Frith, 2008). According to this evolutionary perspective, the ability to make an immediate judgement about certain characteristics of others has served an important adaptive purpose throughout the evolution of humankind, providing some individuals with survival and reproductive advantages (e.g., McArthur & Baron, 1983; Oosterhof & Todorov, 2008; Sell et al., 2008; Short, Mondloch, McCormick, Carré, Ma, Fu, & Lee, 2012). To enhance survival capacity, it is important for humans to possess an ability to detect threats to resources and safety that are posed by other humans and to recognize opportunities for reproduction (McArthur & Baron, 1983; Oosterhof & Todorov, 2008; Sell et al., 2008; Short et al., 2012). As such, to enhance one's own evolutionary fitness (i.e., the likelihood of passing on one's genes to future generations), and the fitness of the group one belongs to, immediate detection of certain characteristics in others is crucial for the accurate detection of threat-relevant characteristics (Damasio, 1994; Frith & Frith, 2008; Schaller, 2008). Furthermore, peoples' accuracy in perceptions of threat, specifically, has been explained in terms of evolutionary survival techniques. For example, primates were able to conserve resources and avoid life-threatening attacks when losing in hostile situations by conveying dominant signals and submissive signals, respectively (de Waal, 2007). The success of these strategies over evolutionary time has likely helped to fine-tune detection mechanisms for cues to vulnerability and dominance (e.g., walking style, body language, gesturing; Grayson & Stein, 1981; Murzynski & Degelman, 1996; Richards et al., 1991).

Gender Differences in Perception. If immediate social perception is an evolutionarily-

evolved adaptation, all humans, regardless of gender, should demonstrate considerable accuracy in their judgements of others. Research on gender differences in interpersonal perception, however, has resulted in mixed findings. Kohn and Fiedler (1961) investigated the effect of gender on participants' perceptions of significant people in their lives (e.g., mother, favourite teacher, best friend). To do so, participants were asked to rate the significant people in their lives on 20 evaluative 6-point bipolar scales (e.g., friendly – unfriendly). The results of Kohn and Fiedler (1961) revealed consistent gender differences. Specifically, females tended to rate themselves as being more similar to their significant others, who they consistently rated in a more favourable light (Kohn & Fiedler, 1961). The authors posited that this finding may suggest a greater susceptibility of women to stereotyping when making interpersonal judgements and that women may evaluate people more rapidly than men, without considering all of their characteristics (Kohn & Fiedler, 1961). Unfortunately, this study did not examine actual differences in *accuracy* between males and females. Furthermore, the results from this study are limited to attempting to explain gender differences in the perception of people who are significant in individuals' lives. Thus, it is possible that such gender effects only exist in the evaluation of people who are known well by the individual. Additionally, the results from this study may not extend to automatic social judgements or the use of non-verbal cues in forming first impressions of strangers.

Similarly, Capetillo, Smith, and Kirby (2008) examined potential differences between males and females in perceptions of other people. Participants were asked to read vignettes and imagine themselves in the social situation described therein in which they needed to accomplish a task with an unhelpful partner. The only differences between the vignettes were the gender (i.e., male or female) and attractiveness (i.e., 'attractive' or 'unattractive') of that partner

(Capetillo et al., 2008). The results provided support for the conceptualization of person perception as a universal adaptation, revealing no differences across gender in terms of their perceptions of the unhelpful partner (i.e., they were equally unwilling to interact with that partner in the future and equally negative in their evaluations). The only gender difference identified was that male participants tended to perceive attractive women as much more attractive than did female participants (Capetillo et al., 2008). Unfortunately, while this study provides support for the perspective that person perception is a universal adaptation, the potential interpretations that can be taken from the results are limited in scope (i.e., only negative perceptions from negative vignettes). Furthermore, the experimental design limits the conclusions that can be drawn from the results. That is, since fictional vignettes were used instead of real targets, comparisons cannot be made between participant's ratings and true ratings of target violence (i.e., self-report or otherwise). Therefore, one cannot draw any conclusions about potential gender differences in the *accuracy* of perceptions of other people.

Chan, Rogers, Parisotto, and Biesanz (2010), on the other hand, did examine gender differences in accuracy in the formation of first impressions. In their study, they examined individuals' abilities to accurately perceive information about targets' personalities, both distinctively and normatively. Distinctive accuracy concerns the ability to perceive aspects of an individual that are unique to them (i.e., how one person is different from the average person), whereas normative accuracy concerns a general understanding of how traits vary among targets (i.e., how ones' impressions of targets relate to an average person's level on given traits; Chan et al., 2010). To assess these different types of accuracy, participants either watched videos of seven females answering "'getting-to-know-you' questions" (Chan et al., 2010, p. 118) or engaged in an unstructured interaction with each other participant for three minutes. Following

each stimulus, participants rated target individuals (i.e., females in the video or partners in the unstructured interactions) on their personality traits. The results indicated a gender difference in normative accuracy. That is, females achieved greater normative accuracy than men, indicating that women's impressions of others relate more strongly to the average person than do men's impressions (i.e., more accurate generalized impressions across targets; Chan et al., 2010). This result seems to counter the notion that person perception is an adaptation that evolved to enhance survival and reproduction capacity in both males and females over time. However, there was no gender difference in distinctive accuracy – participants achieved high levels of distinctive accuracy in their perceptions, regardless of gender. Thus, when it comes to impressions about the unique characteristics of a person, males and females may be equally accurate (Chan et al., 2010). This suggests that perhaps the ability to perceive specific information about a given person in a certain situation has provided an evolutionary advantage to humans, regardless of gender.

Similarly, Stillman et al. (2010) found that males and females were equally accurate in distinguishing individuals with a violent propensity from those without based on first impressions. They showed participants photographs of sex offenders' faces, each for two seconds, and had them rate how violent they perceived the individual in the photo to be (Stillman et al., 2010). For the purpose of their study, the measure of violence was whether or not violence was included in the targets' offences. Violent sex offences were classified as those including an additional physical violence component above and beyond the sexual offence, while non-violent sex offences were classified as those which did not¹ (Stillman et al., 2010). Their finding that males and females are both able to accurately perceive a propensity for violence just by quickly

¹ Violent and non-violent classifications for Stillman et al.'s (2010) study will be described in greater detail below.

looking at peoples' faces provides support for the assumed universality of person perception as an adaptation, but also speaks more specifically to an adaptively-tuned mechanism for threat detection. Furthermore, as the stimuli were only presented for two seconds each, the results support the conceptualization of this mechanism as an automatic process that likely served to help humans survive throughout our evolutionary history. While Stillman et al. (2010) did not find that one gender was any more or less accurate than the other, they did find that women tended to rate the men in both categories (i.e., violent and non-violent) as more violent overall. This phenomenon may also have emerged as a result of evolutionary pressures acting to develop specific adaptations to avoid or reduce costs to oneself. This will be explored further in the following section.

Evolutionary Frameworks for Variations in Threat Detection. If the ability to detect threat in others is, in fact, an evolutionarily-evolved ability, one may question why mistakes are sometimes made in this perception or why there may be variation within accurate responding, as in the results of Stillman et al.'s (2010) study. However, Haselton, Buss, and DeKay (1998) proposed that even variations in the perception of others result from evolutionary processes. According to *Error Management Theory* (EMT), certain decision-making biases in the perception of threat in others have evolved as adaptive mechanisms. According to this perspective, throughout evolutionary history, humans were recurrently faced with making decisions under uncertain circumstances. These decisions often involved two or more options that were associated with differing levels of cost (Haselton & Buss, 2000; Haselton et al., 1998). As a result, selection processes fashioned a decision-making bias by which people tend to choose the decision that is associated with the least cost to them, thereby enhancing their chances of survival (Green & Swets, 1966; Haselton & Buss, 2000; Haselton et al., 1998). One simplified

example, for illustrative purposes, may involve a woman who notices a strange man walking toward her when she is alone. As the man is a stranger to her, she is uncertain about whether this man is threatening and must decide if she needs to avoid him to protect herself. In this case, deciding that the man poses a threat and avoiding him is the decision associated with less costly consequences (i.e., losing a potential opportunity for reproduction versus being attacked), which may or may not be the accurate decision. Unfortunately, this perspective does not account for the fact that humans are often quite accurate in their judgements of others (e.g., Ambady & Rosenthal, 1992; Ambady et al., 2006; Funder, 1987; Kenny & Albright, 1987; Stillman et al., 2010). However, it may go some way to explaining certain variations in accurate responding, such as differences in degree of accurate judgements. Specifically, this potential adaptive bias may help to explain the finding by Stillman and colleagues (2010) that women, while no less accurate than men in perceiving violent propensity, tended to judge the targets as more violent overall.

While EMT (Haselton et al., 1998) provides an evolutionary explanation for why people may be biased toward making errors in threat detection, particularly women, Porter and ten Brinke (2009) proposed an evolutionary theory that may more specifically explain why people rely on misleading cues in judgements of threat. The basis of this theory, called *Dangerous Decision Theory* (DDT), is that the ability to rapidly perceive threats by looking only at faces evolved as a survival mechanism but this ability is not as adaptive in the modern day as it has been in our evolutionary past. This immediate perception of non-verbal cues creates lasting impressions which may, in the modern day, lead people to make ‘dangerous decisions’ that could be harmful to others (e.g., deciding in court that a violent man is innocent because of lasting first impressions based on misleading cues, such as facial attractiveness; Porter & ten Brinke, 2009).

This perspective may go some way to explaining why individuals often rely on misleading cues² to inform their decisions about other people. However, like EMT, it does not account for the fact that individuals remain surprisingly accurate in their perceptions of the characteristics of other people (e.g., sexual orientation, aggression, vulnerability, violent propensity; Ambady et al., 1999; Carré et al., 2009; Stillman et al., 2010; Valentova et al., 2011).

Conversely, Duntley and Shackelford (2008) apply an evolution-centered perspective specifically to explain the adaptation of the ability to accurately detect threat in others. They believe that a threat-detection adaptation is the result of an antagonistic type of co-evolution between criminals and victims of crime. Duntley and Shackelford (2008) explain that criminal behaviour has served to benefit criminals at the cost of victims, which has resulted in a co-evolutionary arms race. Being a victim of violent crime has numerous negative effects on one's evolutionary fitness, which create selection pressures on victims for the evolution of defense adaptations (e.g., a murder victim is unable to protect children and loses future chances of reproduction; Duntley & Shackelford, 2008). Thus, as a result of the adaptation of violence and hostility in humans, the counter-adaptation of an ability to detect cues to threat was necessitated. This counter-adaptation similarly created new selection pressures on perpetrators of crime, causing the evolution of further adaptations for the exploitation of victims. This antagonistic arms race has continued as such throughout evolutionary time, developing the ability in humans to accurately perceive threats in others (Duntley & Shackelford, 2008). Figure 1 provides a visual summary of this co-evolutionary push-and-pull.

Evolutionary perspectives and frameworks for explaining humans' abilities and errors in perceiving the characteristics of others have provided a good basis for the advancement

² Misleading cues will be discussed in further detail below.

Figure 1. An illustration of the antagonistic arms race resulting from selection pressures between victims and perpetrators adapted from Duntley, J. D., & Shackelford, T. K. (2008). *Evolutionary forensic psychology: Darwinian foundations of crime and law* (pp. 201-229). New York, NY: Oxford University Press.

of research in person perception. For example, evolutionary frameworks that account for certain errors made in person perception, such as the use of misleading cues, have led to a wealth of research on perceptions of others and what cues may erroneously sway individuals' decision-making (e.g., Curno, 2015; Stillman et al., 2010). Moreover, interpreting these perceptive abilities as adaptive mechanisms has led researchers to formulate and test numerous hypotheses consistent with this perspective, leading to a greater understanding of humans' success in judging the characteristics of others from very limited information. For example, researchers have begun investigating the role of distinct personality traits in the ability to successfully perceive the

characteristics of others, finding that certain personality traits do, in fact, enhance ones' abilities in such perception (e.g., Black, Woodworth, & Porter, 2014; Ritchie, 2016). These results suggest that one aspect of a person that may lead to differential abilities to accurately perceive the characteristics of others may be his or her personality. Thus, in order to fully comprehend variations in person perception, an overview of certain personalities that relate to differential abilities to judge the characteristics of others is necessary.

Personality

The Psychopathic Personality. The psychopathic personality is characterized by a combination of maladaptive interpersonal, affective, behavioural, and antisocial personality traits, including glibness, callousness, impulsivity, manipulation, and a lack of empathy and remorse (Cleckley, 1941; Hare, 1991, 2003). Measures of psychopathy have historically been organized in terms of two central Factors of the personality. Factor 1 accounts for the interpersonal and affective aspects of psychopathy and Factor 2 accounts for the traits related to unstable lifestyle and antisocial behaviour (Hare, 1991). Recently, however, these two facets have been subdivided to create four facets of psychopathy: interpersonal, affective, lifestyle, and antisocial (see Figure 2 for a summary of this factor structure). Many self-report measures of psychopathy have been developed to represent this factor structure for the assessment of psychopathy in community and student samples. For instance, the Self-Report Psychopathy Scale (SRP-4; Paulhus, Neumann, & Hare, 2016) is a self-report version of the Hare Psychopathy Checklist-Revised (PCL-R; Hare, 2003), a measure of psychopathy commonly used in clinical settings. Like the PCL-R, this measure produces four separate facet scores that compose the two main factors of psychopathy. The four facets of the SRP-4 are interpersonal manipulation (IM), callous affect (CA), erratic lifestyle (EL), and antisocial behaviour (AB).

Figure 2. A summary of the four-facet conceptualization of psychopathy.

The sum of the first two facets provides a score for Factor 1 psychopathy and the sum of the latter two facets provides a score for Factor 2 psychopathy (Paulhus et al., 2016).

The four factor conceptualization of psychopathy has been replicated within various groups, such as offender, youth, forensic psychiatric, and general populations (Babiak, Neumann, & Hare, 2010; Kosson et al., 2013; Neumann, Kosson, & Salekin, 2007). However, some researchers have purported that alternative models may better account for the central features of psychopathy (e.g., Cooke & Michie, 2001; Patrick, 2010). Thus, other self-report measures exist that are based on alternative factor structures. These tend to maintain some core themes of psychopathy (e.g., lack of empathy) but differ in conceptualizations of the most important features (Goodwin, 2014). The Triarchic Psychopathy Measure (TriPM; Patrick, 2010), for example, proposes a three-factor model for assessing psychopathy. This model was proposed, in part, because it has been argued that the PCL-R does not capture the boldness element of psychopathy. The three factors captured in this model are three distinct personality

constructs considered to be related to psychopathy: boldness, meanness, and disinhibition (Patrick, 2010). However, much debate exists in the literature concerning the usefulness of the boldness facet in the assessment of psychopathy. Many academics maintain that it is a crucial aspect of the psychopathic personality (e.g., Lilienfeld, Patrick, Benning, Berg, Sellbom, & Edens, 2012; Lilienfeld et al., 2016; Patrick, Venables, & Drislane, 2013). Others have concluded otherwise, claiming that boldness is either unrelated to psychopathy or non-essential to its measurement (e.g., Crego & Widiger, 2014; Lynam & Miller, 2012; Miller & Lynam, 2012; Neumann, Uzieblo, Crombez, & Hare, 2013). Thus, it remains unclear which factor structure best captures the essential elements of psychopathy, as both three and four factor models have been empirically supported (e.g., Cooke & Michie, 2001; Cooke, Michie, Hart, & Clark, 2005; Edens, Marcus, Lilienfeld, & Poythress, 2006; Hare, 2003).

Psychopathic individuals tend to live their lives in a selfish manner, acting to serve their own interests through callous manipulation, deception, and exploitation, regardless of any potential detriment their actions may inflict on others (e.g., Book et al., 2007; Book et al., 2013; Jones & Paulhus, 2011). Thus, not surprisingly, these individuals are often very successful in using their superficial charm and expert manipulation and deception skills to exploit victims (e.g., Hare, 2001). Therefore, it is a logical assumption that the personality traits associated with the psychopathic personality may provide an advantage to individuals in the perception of others. In fact, research has demonstrated that psychopathic individuals are better able to identify vulnerability in others than non-psychopathic individuals (e.g., Book et al., 2007; Ritchie, 2016). However, research has not yet examined whether this enhanced perception extends to the judgement of characteristics in others that do not relate to vulnerability.

HEXACO Model of Personality. As it has been demonstrated that psychopathic individuals may have an enhanced ability to accurately detect the characteristics of others, it may also be possible for general personality traits to be associated with such accuracy. The original ‘Big Five’ Factors of general personality (i.e., Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Intellect/Imagination; Goldberg, 1990, 1993), discovered through English-language lexical studies, were the first to lead to a consensus among many researchers regarding the structure of personality. The Five Factor Model (FFM) incorporated the lexical dimensions of the Big Five into a very similar structure, including Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience (McCrae & Costa, 1987). The constructs in this model of personality were popularly used for decades of personality research as reliable and valid markers of variations in personality. However, subsequent comprehensive lexical studies of personality resulted in the identification of six major personality dimensions and a shift in the conceptualization of personality was proposed (Ashton et al., 2004b; Lee & Ashton, 2004).

The HEXACO Personality Inventory (HEXACO-PI; Lee & Ashton, 2004) and its short-form, the HEXACO-60 (Ashton & Lee, 2009), are questionnaires based on the six-dimensional model that resulted from numerous lexical studies on the structure of personality – the HEXACO (Ashton, Lee, & Goldberg, 2004a; Ashton, Lee, & de Vries, 2014; Ashton et al., 2004b; Lee & Ashton, 2004, 2008). This model is similar to its predecessors, the Big Five and the FFM, in many ways as it is a refined extension of these models of personality. The title of the HEXACO is an acronym for the personality dimensions contained within: (H) Honesty-Humility, (E) Emotionality, (X) eXtraversion, (A) Agreeableness, (C) Conscientiousness, and (O) Openness to Experience. The same six dimensions have been replicated in lexical studies cross-culturally,

emerging across at least 12 different languages (Ashton & Lee, 2007), and some scholars posit that this model of personality more accurately represents variations in personality (Ashton et al., 2014). Each of these six dimensions has four associated facets that are captured by the HEXACO-PI (Ashton & Lee, 2007). Table 1 contains a list of the six dimensions of the HEXACO model of personality and their associated facets.

Table 1

HEXACO Dimensions and Corresponding HEXACO-PI Facets

HEXACO Dimension	HEXACO-PI Facets
Honesty-Humility	Sincerity Fairness Greed-Avoidance Modesty
Emotionality	Fearfulness Anxiety Dependence Sentimentality
Extraversion	Expressiveness Social Boldness Sociability Liveliness
Agreeableness	Forgiveness Gentleness Flexibility Patience
Conscientiousness	Organization Diligence Perfectionism Prudence
Openness to Experience	Aesthetic Appreciation Inquisitiveness Creativity Unconventionality

The HEXACO-60 was designed to provide a short measure for the assessment of the personality constructs identified in the HEXACO model of personality for use when administration time is limited (Ashton & Lee, 2009). Ten items were chosen from each of the six scales of the HEXACO-PI (Ashton & Lee, 2008; Lee & Ashton, 2004) to ensure all content contained within each dimension was adequately captured. This short-form has demonstrated good reliability and validity and is recommended for the measurement of personality when time constraints exist (Ashton & Lee, 2009).

Certain dimensions of the HEXACO model of personality have been associated with features of psychopathy, namely low honesty-humility, low conscientiousness, and low agreeableness (e.g., Ashton, Lee, & Son, 2000; Ashton & Lee, 2008; Lee & Ashton, 2005). As psychopathy may be related to an enhanced generalized ability to accurately perceive characteristics of others, those personality dimensions of the HEXACO that are associated with psychopathy may also increase such accuracy. In fact, it has been found that individuals with lower scores on the honesty-humility facet are more accurate in perceptions of vulnerability to violence (Ritchie, 2016). However, it is unclear whether this perception accuracy extends to the general perception of others and, thus, to the detection of threat in others. As such, an exploration of which HEXACO dimensions of personality, if any, contribute to accuracy in threat detection is warranted.

Research on Immediate Interpersonal Judgement

Academics have long discussed the adaptive ability of humans to assess characteristics of others, such as their personality traits, emotional states, social relations, and motivations, simply from brief exposures to non-verbal cues (Allport, 1937; Ambady, Bernieri, & Richeson, 2000; Ambady, Krabbenhoft, & Hogan, 2006; Goffman, 1979). Non-verbal cues are classified as static,

dynamic, or paralinguistic (Degroot & Gooty, 2009). Static cues are those characteristics of an individual that cannot be easily changed, such as physical appearance or demographic characteristics. Dynamic cues, on the other hand, can readily change, such as head movement or facial expression. Paralinguistic cues are those non-verbal cues that relate to characteristics of an individual's voice, such as how quickly or how loudly they speak (Degroot & Gooty, 2009). Any number of these non-verbal cues can convey information about an individual's personality.

In situations where inadequate information is available to form well-reasoned responses, individuals rely upon a highly automatic, implicit, fast, inflexible, and efficient form of processing. Under such circumstances, humans quickly and subconsciously process the situation, considering only minimal information and relying on existing knowledge and expectations (e.g., Bohl & Bos, 2012; Frith & Frith, 2008). While these immediate trait judgements of others are based on seemingly shallow information, non-verbal expressions may actually be a more accurate depiction of a person's characteristics than anything they communicate verbally (Ambady et al., 2006). This is true for a number of reasons that uniquely relate to the expression of non-verbal behaviours (DePaulo, 1992). For instance, it is much more difficult for a person to control instantaneous non-verbal behaviour than it is to control what he or she says. This fact often leads to 'leakages' of true intentions or feelings. Another aspect of this phenomenon is that one's own non-verbal behaviours are not as apparent to oneself as they are to others, which makes self-monitoring of these cues difficult (Ambady et al., 2006). It is also much quicker to express non-verbal cues than to express oneself verbally. A person may, for example, exhibit a facial expression indicative of a certain emotional state before he or she has the chance to stifle that expression. In this sense, non-verbal expressions may actually provide more honest

manifestations of a person's characteristics, traits, and intentions than the things they say (Ambady et al., 2006).

One body of research has been dedicated to investigating the role of the automatic processing of non-verbal cues by rapidly presenting stimuli to participants and having them rate targets on certain characteristics or motivations (e.g., Ambady & Rosenthal, 1992; Willis & Todorov, 2006). The stimuli, which may be dynamic or static, are often very short (i.e., a number of milliseconds or a couple of seconds – never longer than 5 minutes), may not involve any interpersonal interaction (Albright, Kenny, & Malloy, 1988; Ambady et al., 2006), and emphasize salient information (Ambady et al., 2006). From brief stimuli such as these, individuals form automatic assessments based on cues perceived from others' facial expressions, gestures, voice, and body movements (Ambady et al., 2000) and are, in fact, quite accurate in doing so (e.g., Ambady & Rosenthal, 1992; Ambady et al., 2006; Funder, 1987; Kenny & Albright, 1987). These assessments, in turn, influence the way that individuals act in social situations (e.g., Ambady et al., 2006; Jones, 1990).

This area of investigation has demonstrated that a variety of non-verbal cues are processed by individuals to construct rapid evaluations of other people that are surprisingly accurate. For instance, laypeople are able to evaluate aspects of other people such as their sales effectiveness (i.e., of salespeople; Ambady et al., 2006), vulnerability (e.g., Ritchie, 2016), performance on job evaluations (Ambady & Rosenthal, 1993), sexual orientation (Ambady, Hallahan, & Conner, 1999; Valentova, Rieger, Havlicek, Linsenmeier, & Bailey, 2011), and infidelity (Lambert, Mulder, & Fincham, 2014). Of particular interest is that the existing literature has also demonstrated that the ability to quickly and accurately judge target

characteristics extends to characteristics that relate to threat, such as dominance, aggression and violence, and psychopathy.

Perception of Dominance. Schmid Mast and Hall (2004) sought to determine whether people are accurate in judgements of status based on non-verbal cues (i.e., relative dominance and submissiveness). Their participants were presented with candid photographs of pairs of university staff within the same department engaging in a naturalistic interaction. These targets completed self-reports regarding their relative status to their co-worker. In the majority of the photos, the targets were sitting, their full body was visible, and their faces were visible from the side. However, in some photos, a full view of targets' bodies was obstructed by a desk and, in some cases, their faces were not visible (i.e., facing away from the camera; Schmid Mast & Hall, 2004). One sample was asked to rate the status of each member of the dyad. For this group of participants, the targets were cut out of the photographs and pasted on white paper to eliminate context cues, maintaining the distance between the dyad. A second sample was asked to rate the individual targets' status. In this case, one member of the dyad was covered up and the participants were asked to rate the status of the target in relation to their co-worker (i.e., higher or lower hierarchical position in the workplace; Schmid Mast & Hall, 2004). The results demonstrated that, regardless of the gender composition of the target dyad or the gender of the single target, participants were able to accurately perceive differences in status based solely on non-verbal cues. This was equally true for both male and female observers (Schmid Mast & Hall, 2004). However, the procedures used in this study raise questions about the quality of the measures (e.g., cutting and pasting, obstruction of faces or bodies in some photos) and could potentially have affected the results. For example, cutting and pasting targets from photographs could have changed targets' angles, making them appear as if they were leaning in or away more

than they were. Thus, there is a need for future research to replicate these findings using a more controlled methodology. Furthermore, dominance may be only one aspect of assessments regarding status. Thus, for specific interpretations of these findings in relation to males' and females' abilities to perceive dominance from non-verbal cues, they should be interpreted in combination with similar findings from studies specifically designed to examine this ability. Moreover, although these results provide evidence for an ability in humans to perceive status from non-verbal cues, the methodology does not allow for conclusions to be drawn about whether this happens by a rapid decision-making process. That is, because the time that participants were presented with the photographs was not limited, it cannot be determined whether status judgements were made by first impression decisions or by a contemplative decision-making process.

The study conducted by Koppensteiner, Stephan, and Jäscke (2016) presents the opposite dilemma. That is, this study provided support for the fact that individuals rely on non-verbal cues in first impression perceptions of dominance, but did not include a true measure of dominance with which to assess accuracy of such perceptions. In this study, brief video clips of male and female politicians delivering speeches were selected. The motions of the politicians in these videos were then mapped onto stick-figure models in order to eliminate extraneous cues (e.g., clothing, appearance) so that the use of motion in first-impression judgements could be evaluated (Koppensteiner et al., 2016). On average, the length of a video clip was 15 seconds. These clips were presented without audio and participants were asked to rate the degree to which they agreed that the individual in the clip was dominant. The results demonstrated that the expansiveness of a single posture, the expansiveness of motion overall, and the amount of motion, were all related to perceived dominance by participants (Koppensteiner et al., 2016). While these findings do not

speak to the accuracy of first impression judgements of dominance, they do provide support that immediate judgements of others are influenced by simple non-verbal cues – in this case, motion.

Furley and Schweizer (2016), conversely, examined the accuracy of rapid perceptions of dominance by laypeople. The motivation for their study was based on the assertion that, evolutionarily, exuding submissiveness in a life-threatening situation decreased the chances of primates to experience further attacks (de Waal, 2007) and that, conversely, displaying dominance and superiority in such a situation helped primates to conserve resources (Archer, 2006; de Waal, 2007). Thus, to examine whether human beings are able to assess dominance and vulnerability through non-verbal channels of communication, Furley and Schweizer (2016) had participants rate the players in brief video clips of professional basketball games on a variety of characteristics related to status (i.e., dominance, confidence, and pride). These videos were selected to make sure that no obvious cues that may demonstrate such personality traits were blatantly depicted (e.g., hiding one's face in one's hands). The findings demonstrated that people were very accurate in distinguishing leading athletes from trailing athletes. Based solely on the non-verbal behaviour they saw in the video clips, participants accurately identified the leading athletes as more dominant, confident, and proud than the trailing athletes (Furley & Schweizer, 2016).

Rule, Adams Jr., Ambady, and Freeman (2012) also examined the capacity of humans to accurately detect cues relevant to dominance and submissiveness under physical and temporal constraints. In a three-part study, they examined this accuracy in relation to the use of facial cues and in relation to the use of bodily cues. In study one, participants were shown standardized photographs (i.e., cropping, size, grayscale) of Caucasian men displaying neutral facial expressions, as well as posing to express prototypical expressions of dominance (i.e., head tilted

upward and a direct gaze) and of submissiveness (i.e., head tilted downward and an averted gaze). Another condition in this study was the display of altered versions of these photographs that just showed major facial outlines (i.e., low spatial frequency information was removed). A similar methodology was implemented for study two, but with standardized photographs of full bodies displaying such poses. In each condition (i.e., face or body), participants viewed the photographs for either 27 milliseconds (ms), 40 ms, 94 ms, or at a self-paced rate (i.e., unlimited time). The results from this study demonstrated that participants were able to accurately perceive dominance and submissiveness from exposure to stimuli as brief as 40 ms (for face or body). Further, greater exposure time increased perception accuracy, which may be explained by the fact that obvious prototypical cues were blatantly posed by targets (Rule et al., 2012). In unposed photographs, however, greater exposure time has simply increased confidence in judgements, rather than causing changes to the initial assessment (Willis & Todorov, 2006). Accuracy was higher for judgements that were based on faces and for photos in which both high and low frequency information was available (i.e., not just outlines; Rule et al., 2012). However, participants were able to accurately perceive dominance and submissiveness from brief glances (i.e., as little as 1/25th of a second) at just the outline of a face or body displaying prototypical cues. Rule et al. (2012) also concluded from their results that, in the judgement of dominance and submissiveness, humans preferentially rely on facial cues. In sum, according to the existing research in this area, individuals are able to accurately assess other peoples' proclivities toward dominance and submissiveness from very little information. Thus, it is not surprising that, without much information, human beings are also able to accurately assess when others are likely to be aggressive or violent.

Perception of Aggression and Violence. As people are able to discern characteristics such as vulnerability and dominance from very brief exposures to non-verbal stimuli, it is not surprising that they have similarly been quite successful in identifying propensity toward aggression and violence in others. Willis and Todorov (2006) examined the tendency of laypeople to make rapid trait judgements about people by just looking at pictures of their faces. They used standardized facial photographs (e.g., all wore gray T-shirts, no makeup, facial hair, earrings, etc.) of targets displaying neutral facial expressions. Among the trait dimensions to be rated was aggression. However, aggression was not quantified by a measure of true aggressive propensity but, rather, by using mean ratings of the degree to which each target was perceived to be aggressive by a pilot sample (Willis & Todorov, 2006). Thus, this was a measure of consensus on the appearance of aggression rather than accuracy in threat detection. However, ratings under a 100 ms time constraint significantly correlated with ratings made without any time constraints, providing evidence that people do make trait inferences about aggression in as little as 100 ms. Additionally, Willis and Todorov (2006) used photographs of both males and females, providing some evidence that this immediate assessment of threat posed by others occurs regardless of their gender.

Bar, Neta, and Linz (2006) similarly examined how quickly observers form first impression judgements of threat. They used 90 photographs of Caucasian male faces displaying neutral expressions. The photographs were presented in grayscale on a gray background. Bar et al. (2006), like Willis and Todorov (2006), examined consensus between time conditions (i.e., 27 ms, 39 ms, 1700 ms, or no limit) among ratings of the degree to which target faces were believed to belong to threatening people. The results from their analyses suggest that people make decisions about the potential threat of others from exposures as short as 39 ms that are consistent

with decisions made from longer exposures (i.e., 1700 ms and unlimited time). This indicates that humans make perceptions about threat based on whatever little information is available within 39 ms of exposure to another person (Bar et al., 2006). Moreover, Bar et al. (2006) found preliminary evidence that the difference in consistency between 39 ms exposures and shorter exposures may be related to the level of participant awareness of facial information. However, viewing a face for as little as 39 ms increased this awareness by enough that participants were able to form consistent impressions of threat.

Using photographs of varsity hockey players, Carré et al. (2009) conducted two studies in which they did examine the accuracy of observers' ratings of aggressive propensity. The participants' ratings were based on static facial features of targets for whom violence scores had been previously quantified (i.e., using penalty minutes in hockey games; Carré & McCormick, 2008). These scores were used to make comparisons between participants' ratings of violence and actual violent propensity of the targets to determine the accuracy of participant ratings. In study one, the faces were displayed to participants in random order on a black background, each for 1,000 ms, to familiarize observers with the range of faces. They were then presented to participants again in randomized order, each for 2,000 ms, after which observers were asked to rate the face on propensity for violence, dominance, masculinity, trustworthiness, and attractiveness (Carré et al., 2009). In the second study, no familiarization with the faces took place. To direct attention, a central fixation cross was displayed for 500 ms before each of the faces, which were presented for only 39 ms each (Carré et al., 2009). Observers in this study were then asked to estimate the target's propensity toward violence. The results demonstrated that observers are able to accurately estimate targets' propensities toward violence simply based upon very brief glances at neutral photographs of their faces (Carré et al., 2009). The fact that

participants were able to make accurate inferences about target's violent inclinations provides a good demonstration of the adaptive ability of humans to immediately perceive threat in other people. However, the sample sizes for these studies were quite small (i.e., 31 participants in study one and 16 participants in study two). Thus, these findings should be interpreted in relation to findings from other studies that have produced similar results with larger samples of observers.

Stillman et al. (2010) also examined the ability of laypersons to accurately detect a propensity toward violence based on brief exposures to static facial features. They presented participants with facial photographs of registered sex offenders collected from a sex offender registry database. Of the offenders in these photographs (targets), 37 were considered non-violent and 47 were considered violent, based on their publicly available criminal histories. These photographs were displayed to participants, each for two seconds. Based on the rapid exposures to the photographs, participants were asked to rate the targets' propensities for violence on a Likert-scale ranging from 1 (*less likely to be violent*) to 4 (*more likely to be violent*; Stillman et al., 2010). The results demonstrated that participants were accurate in their assessments of violent propensity. That is, estimated likelihoods of violence were significantly related to actual violent histories of the targets. Stillman et al.'s (2010) sample size was much larger than those of Carré et al. (2009), with 97 observers rating the photographs, which may indicate an increased statistical power associated with their findings. However, this sample size is relatively small and, thus, replication of these results is needed.

In sum, research demonstrates that people are quite accurate in their judgements of others' propensity for violence and aggression. Relatedly, it is not surprising that research has also demonstrated an ability of humans to accurately detect psychopathy (Fowler et al., 2009;

Gillen, Bergström, & Forth, 2016), a personality type that has consistently been linked with violent offending (e.g., Leistico, Salekin, DeCoster, & Rogers, 2008; Wormith, Olver, Stevenson, & Girard, 2007).

Perception of Psychopathy. Research has demonstrated that laypeople are able to accurately judge psychopathy based on very brief clips of interviews (Fowler et al., 2009). Fowler and colleagues (2009) demonstrated that laypeople can make accurate judgements of others based on brief excerpts of interviews and that they can do so regardless of presentation modality. In their study, participants were shown 5-second, 10-second, or 20-second audio, video, or combined audio-video clips of interviews with maximum security inmates using the PCL-R. Based on these clips, they were able to accurately assess the inmate's level of psychopathy, particularly when the clips were presented for 5 or 10 seconds and in the video or audio-video modalities. These findings indicate that non-verbal cues are particularly salient in assessments of psychopathy. Further, the strongest relationships between assessors' ratings and actual measures of psychopathy were found under the condition of the shortest exposure time (5 seconds). This finding suggests that characteristics of psychopathy may be more reliably detected when assessors have little information to go on. Fowler and colleagues (2009) posited that this may relate to the superficial charm inherent in psychopaths which, when time permits, may be used by psychopathic individuals to ease any initial misgivings assessors may have.

Using brief interview clips, Gillen et al. (2016) also examined the ability of laypeople to accurately detect psychopathy. In this study, the rater's scores on certain Big Five personality dimensions (consisting of conscientiousness, agreeableness, neuroticism, extraversion, and openness to experience; Costa & McCrae, 1992) were related to increased accuracy. In their study, the targets were female offenders. Thus, specifically, the results demonstrated that

personality traits of participants (i.e., the raters) that relate to social experience and the detection of non-verbal cues were predictive of greater perception accuracy of psychopathic traits in female offenders. Factor 1 traits (i.e., interpersonal and affective characteristics) were better judged by participants with high ratings of agreeableness and low ratings of behavioural-antisocial psychopathic traits and Factor 2 traits (i.e., lifestyle and antisocial characteristics) were better judged by participants with higher ratings of extraversion-sociability. Gillen et al. (2016) posited that these findings might be explained by a broader social experience and an increased social competence of raters with such personality traits. Psychopathic individuals are callous, manipulative, and lack remorse (Cleckley, 1941; Hare, 1991, 2003) and, as such, they easily exploit people for their own gain, regardless of the damage they may inflict (e.g., Book et al., 2007; Book et al., 2013; Jones & Paulhus, 2011). Thus, the ability of people to detect psychopathic traits in others based on non-verbal cues provides support for the existence of an automatic threat-detection mechanism in humans.

Cues to Aggression and Violence

In studies of the innate threat detection abilities of humans, a number of non-verbal cues have been identified that are used by individuals in forming judgements about other people, such as physical and interpersonal attractiveness (e.g., Capetillo et al., 2008; Castellow, Wuensch, & Moore, 1990; Fowler et al., 2009) and inadvertent non-verbal behaviours (i.e., ‘leakages’ of true intentions or feelings; e.g., Furley & Schweizer, 2016; Pretsch et al., 2013; Richards et al., 1991). It has been repeatedly demonstrated that cues detected from faces greatly inform judgements about the traits and propensities of others. Many of these cues are valid indicators of aggressive and violent propensities, informing accurate decisions regarding threats posed by other people.

However, there are some non-verbal facial cues that may be misleading or irrelevant which people also tend to rely on to inform erroneous decisions about threat posed by others.

Testosterone and Facial-Width-to-Height-Ratio. The surprising accuracy of individuals' trait judgements of aggression and violence based simply on facial features may be related to the sexual dimorphism of facial features. That is, relative accuracy in the prediction of traits such as trustworthiness, cooperation, and aggression may be related to the expression of male sex hormones, such as testosterone. Testosterone is a hormone that has been found to be related to both facial growth (e.g., Verdonck, Gaethofs, Carels, & de Zegher, 1999; Penton-Voak & Chen, 2004; Carré & McCormick, 2008) and aggression (e.g., Salvador, Suay, Martinez-Sanchis, Simon, & Brain, 1999; Carré & McCormick, 2008, 2011). In humans, testosterone level is linked to degree of facial masculinity (Penton-Voak & Chen, 2004). Specifically, in addition to other sexually dimorphic characteristics, such as facial hair, a younger appearance, and a heavy brow for males (Stillman et al., 2010), testosterone is linked to an individual's facial width-to-height ratio (FWHR). FWHR provides a measure of masculinity that is independent of body size or strength and is associated with the testosterone level of the individual. That is, a higher level of circulating testosterone leads to a larger FWHR (Verdonck et al., 1999; Carré & McCormick, 2008; Carré et al., 2009; Carré, Murphy, & Hariri, 2013; Short et al., 2012; Stillman et al., 2010; Stirrat & Perrett, 2010). FWHR is calculated from facial photographs by measuring the distance across the cheekbones and dividing that value by the distance between the upper lip and the mid-brow (i.e., the upper facial height; e.g., Carré & McCormick, 2008; Carré et al., 2009; Carré, Morrissey, Mondloch, & McCormick, 2010; Carré et al., 2013; Short et al., 2012; Stirrat & Perrett, 2010). It is a non-verbal cue that not only has repeatedly been proven to have utility in the assessment of aggression (e.g., Carré & McCormick, 2008; Carré et al., 2009;

Lefevre & Lewis, 2014) but is also relied upon in these assessments by adults and children alike (Short et al., 2012; Stirrat & Perrett, 2010).

FWHR has been correlated with traits and actions such as cheating (Haselhuhn & Wong, 2012), aggression (e.g., Carré & McCormick, 2008; Carré et al., 2009; Lefevre & Lewis, 2014), and lack of trustworthiness (Stirrat & Perrett, 2010). Many researchers have focused their attention on the relationship between aggression and FWHR. Carré and McCormick (2008), for instance, conducted three studies to examine the relationship between FWHR and self-reported dominance and naturalistic aggression. In the first study, university students had neutral facial photographs taken (to be measured for FWHR), completed self-report measures of dominance, and engaged in a Point Subtraction Aggression Paradigm with a computerized partner who they thought was a same-sex partner on the task (Carré & McCormick, 2008). Throughout this task, participants were given the options to either earn rewards by clicking option 1 one hundred times, subtract rewards from their partner by clicking option 2 ten times (resulting in no change to their own rewards), or protect themselves from point subtraction by clicking option 3 ten times. They were also advised that their partner was able to steal their points, adding to their own rewards (Carré & McCormick, 2008). Evidently, option number two was the measure of reactive aggression for study one. In this study, Carré and McCormick (2008) found that men had larger FWHRs, rated higher on measures of trait dominance, and engaged in more reactive aggression than women. In studies two and three, photographs of professional and varsity hockey players' faces were measured for FWHR and aggression was measured in terms of each player's accumulated penalty minutes per number of games in the hockey season (Carré & McCormick, 2008). Again, differences in FWHR were positively related to differences in accrued penalty minutes. It was concluded, thus, that FWHR may be an accurate signal of a propensity toward

aggression in men. Moreover, these findings provide support for the evolutionary perspective that selection pressures have allowed the perceptual systems of humans to develop the ability to detect cues indicative of aggression in others (Carré & McCormick, 2008).

In a two-part follow-up study conducted by Carré et al. (2009), described in greater detail above, the same hockey-player photographs were used to examine the accuracy of observers' ratings of aggressive propensity based on FWHR. As stated previously, observers were able to accurately estimate targets' propensities toward aggression based on very brief glances at these neutral facial photographs. In addition to this finding, across both studies, faces with higher FWHRs were reliably rated as more aggressive by participants (Carré et al., 2009). Moreover, there was a strong association found between estimated and actual aggression and between estimations of aggression in study one (in which exposure to the faces was 2000 ms each, following a 1000 ms familiarization exposure) and in study two (in which exposure to the faces was reduced to 39 ms). The results from these studies suggest that FWHR is a stable cue that may be used to help inform accurate estimations of individuals' propensities toward violence, even under very strict time constraints (Carré et al., 2009).

Similar results have been found linking FWHR and assessments of aggressive behaviour across race, age (Short et al., 2012), and gender (Lefevre & Lewis, 2014). Short et al. (2012) studied the relationship between the FWHR of photographed own- and other-race targets (i.e., Chinese and Caucasian men) and participants' ratings of aggression, and whether this relationship exists not only for adults' ratings but also for those of 8-year-old children (i.e., Chinese and Caucasian adults and children). The results from this study demonstrated that estimates of aggressive propensity were both strongly and positively correlated with targets' FWHRs and that neither age group demonstrated an other-race bias. As familiarity and

experience with a facial category was unnecessary for significant correlations between aggression ratings and FWHR, Short et al.'s (2012) findings provide support for the proposition that humans are attuned to FWHR in assessments of aggression as a result of evolutionary selection pressures. Moreover, Lefevre and Lewis (2014), using facial photographs manipulated to create varying FWHRs, discovered that this relationship exists even when the targets are female faces. These findings support the conceptualization of an adaptation for threat detection in the form of enhanced perceptual abilities in humans. However, other research indicates that, in addition to attending to valid cues to aggression and violence, people also rely on some misleading and irrelevant cues to inform their assessments.

Valid, Poor, and Misleading Facial Cues to Violence. A study conducted by Stillman et al. (2010) found that participants relied upon a variety of facial cues in formulating their assessments about the violent propensities of the sex offenders they were asked to rate. They classified these into valid, poor, and misleading cues, depending on correlations between the cues and both perceived likelihood of violence (PLV) and actual violent history. Valid cues were those that were related to both a PLV and an actual violent history. Valid cues to violence tended to be those that were indicative of high levels of masculinity: general facial masculinity, a heavier brow, and apparent physical strength. This finding is consistent with the existing research linking aggression and male sex hormones, such as testosterone (e.g., Salvador et al., 1999; Carré & McCormick, 2008). Also consistent with such research is the finding that older age was a valid cue to less violence, as testosterone levels decrease with age (Swerdlow & Wang, 1993). Poor cues were those that they found not to relate to either PLV or violent history, such as smiling, facial hair, and expressions of surprise, fear, or sadness. Attractiveness was also found to be a poor cue (Stillman et al., 2010). However, this seems to contradict the findings of many

other researchers that physically attractive individuals are generally perceived in a more positive light (e.g., Capetillo et al., 2008, Castellow et al., 1990; Curno, 2015; Esses & Webster, 1988; Langlois, Kalakanis, Rubenstein, Larson, Hallam, & Smoot, 2000). Misleading cues were those that were related to estimations of violent propensity but were not related to actual violent history. Facial expressions of disgust and anger commonly influenced mistaken ratings of a higher likelihood for violence, while expressions of happiness, as well as good grooming, influenced mistaken ratings of lower likelihoods for violence (Stillman et al., 2010).

The perception of a mental disorder may be another cue that leads to accurate assumptions of a proclivity toward violence. Research suggests that mentally disordered individuals are significantly more likely than people without mental disorders to be violent (e.g., Arseneault, Moffitt, Caspi, Taylor, & Silva, 2000; Monahan, 1992; Van Dorn, Volavka, & Johnson, 2011). Moreover, surveys have demonstrated that laypeople tend to associate mental disorder with a higher likelihood to engage in violence (e.g., Field Institute, 1984). The belief in a relationship between mental disorder and violence is widespread, as is evidenced by legal policies that target the mentally disordered as a group, as well as informal interactions with individuals who are perceived to be mentally disordered (Monahan, 1992). Thus, if an individual is perceived to be mentally disordered, it seems likely that individuals may rely on these visual cues to inform decisions about the risk that individual poses. In which case, the appearance of being mentally disordered would be a valid non-verbal cue to violence. Based on the existing literature, Table 2 presents a summary of non-verbal cues, as well as their respective validity in violence assessments, that may be used by individuals in assessments of violence in others.

Table 2

Valid, Poor, and Misleading Facial Cues Used in Judgements of Violence and Threat

Valid	Poor	Misleading
General facial masculinity	Smile	Attractiveness *
Heaviness of Brow	Facial hair	Angry expression
Appearance of physical strength	Surprise expression	Disgust expression
Older age*	Fear expression	Happiness expression *
Appearance of mental disorder	Sad expression	Care in grooming *
FWHR		

Note. * = cues that are used to make judgements of less propensity for violence or threat. FWHR = facial width-to-height ratio.

Together, the existing research provides evidence that individuals are able to process stable, non-verbal facial cues automatically and that they rely on these cues to make judgments about the threat-related traits and proclivities of others. Moreover, in the absence of other cues that may signal information about other people's traits and behavioural propensities, such as body size or gait, automatic judgements about interpersonal threat based on assessments of faces demonstrating neutral expressions remain quite accurate (e.g., Carré & McCormick, 2008; Carré et al., 2009; Stillman et al., 2010). This is true for both adults and children regardless of the race (Short et al., 2012) or gender (Lefevre & Lewis, 2014) of the individual being assessed.

The Ability of Psychopathic Individuals to Detect Non-Verbal Cues

The inherent nature of psychopathy allows psychopathic individuals to identify potential victims as targets for exploitation (Hare, 2003; Wheeler et al., 2009) and has allowed them to successfully take advantage of unsuspecting others for personal gain (e.g., sexual or material gain; Harpending & Sobus, 1987 Krupp, Sewall, Lalumière, Sherriff, & Harris, 2012). For this

reason, it has been argued that psychopathy is an evolutionarily evolved strategy that has aided psychopathic individuals in gaining resources and achieving reproductive success (Glenn, Kurzban, & Raine, 2011; Krupp et al., 2012) rather than a mental disorder, as many scholars believe (e.g., Blair, 2010; Corr, 2010; Leedom & Almas, 2012; Newman, Curtin, Bertsch, & Baskin-Sommers, 2010). That is, it has been posited that the characteristics of psychopathy (e.g., promiscuity, coercion, superficial charm) have adaptively evolved as a result of genetic selection, environmental shifts, or a combination of the two (e.g., Glenn et al., 2011). As such, psychopathic individuals have been referred to as ‘social predators’ (Book et al., 2007; Hare, 1991, 2001; Mealey, 1995; Wheeler et al., 2009) who are quite successful in the exploitation of others through the use of deception, manipulation, and charm (Hare, 2001).

According to Frank (1988), the success of social predators, such as psychopathic individuals, likely relates to an increased ability of these people to detect non-verbal cues that signal vulnerability. In fact, research evidence suggests that psychopathic individuals may be more attuned to inadvertent nonverbal cues that inform inherent evaluations of vulnerability than non-psychopathic individuals (Book et al., 2013; Book et al., 2007; Camilleri et al., 2010; Wheeler et al., 2009). For instance, a distinct victim gait has been identified (i.e., a specific victim walking style; Grayson & Stein, 1981; Murzynski & Degelman, 1996) which psychopathic individuals have been particularly adept at reading (Book et al., 2013; Ritchie, 2016). For instance, Book et al. (2013) examined the ability of federal penitentiary inmates with high levels of psychopathic traits to accurately evaluate target vulnerability based only on gait. The inmates rated video clips of twelve individuals who were unknowingly filmed walking (Book et al., 2013). The results demonstrated that inmates who scored higher on Factor 1 interpersonal and affective psychopathic traits, as measured by the PCL-R, were more accurate

in their assessments of victim vulnerability and were more likely to explain the reasoning for their decisions by referencing the targets' gait characteristics (Book et al., 2013).

Perhaps this enhanced ability of psychopathic individuals to read non-verbal cues in others extends beyond only perceptions of vulnerability. If psychopathic individuals are simply more attuned to non-verbal cues, they should also be more accurate than non-psychopathic individuals in judgements of other traits and proclivities. If this ability is not a widely-tuned mechanism for enhanced detection of non-verbal traits, it is also plausible that an increased ability to detect vulnerable targets necessitates an increased ability to detect threat. That is, in choosing an appropriate target for exploitation, perhaps psychopathic individuals 'weed-out' those people who are more likely to harm them. Thus, by this logic, psychopathic individuals may be more accurate in using non-verbal cues for the rapid detection of threat in other people. As such, an evaluation of this ability, as well as which conceptual facets of psychopathy (i.e., IPM, CA, EL, AB, boldness, meanness, disinhibition) relate to such accuracy, is warranted.

Moreover, if it is true that psychopathic individuals have an enhanced generalized ability to perceive and accurately interpret non-verbal cues, characteristics related to psychopathy such as low honesty-humility, low conscientiousness, and low agreeableness (e.g., Ashton et al., 2000; Lee & Ashton, 2005; Lee, Ashton, Wiltshire, Bourdage, Visser, & Galucci, 2013), as measured by the HEXACO model of personality, may also play a part in the accurate evaluation of others' personality traits. Evidence exists that individuals scoring lower on the honesty-humility facet are, in fact, more accurate in perceptions of vulnerability to violence (Ritchie, 2016). Moreover, research has indicated that the personality characteristics of individuals affect the way in which they rate the characteristics of others. For instance, individuals with higher levels of agreeableness, including traits such as altruism and tender-mindedness, often assign lower scores

on the PCL-R (Miller, Rufino, Boccaccini, Jackson, & Murrie, 2011) when assessing offenders. Thus, personality characteristics of perceivers may affect the accuracy of their judgements by influencing their patterns of responding. As such, an exploration of the relationship between the HEXACO dimensions and an ability to accurately detect violence in others is warranted.

Current Study

The goal of the current study was to replicate and extend the research findings of Stillman et al. (2010) using the same facial photographs, a similar methodology, and a larger sample size. They tested the ability of laypeople to detect a propensity toward violence in 84 sex offenders. The stimuli for this study were digital images of the faces of these targets, which were collected from a public sex offender database in America. Due to the existence of strong racial biases regarding threat posed by ethnic minorities, only images of white offenders were included (Maner et al., 2005; Stillman et al., 2010). The targets were classified into violent and non-violent categories based on their publicly-available criminal records. The majority of sex offenders in the violent group had committed aggressive or violent rapes. These were defined as rapes that involved an additional violent component, such as physical assault, using a weapon on the victim, and other sex crimes resulting in the victims' physical injury (Stillman et al., 2010). Non-violent sex offenders were classified, essentially, as those who did not cause physical injury to their victims. These offenders non-violently touched a woman without permission, owned child pornography, and/or engaged in crimes of a statutory nature (Stillman et al., 2010). The photos were then rated by research assistants on a number of non-verbal cues (Stillman et al., 2010). Participants were shown the facial images on a computer, each for two seconds, and asked to rate the targets on their likelihood for violence on a Likert-scale ranging from 1 (*less likely to be violent*) – 4 (*more likely to be violent*). From these ratings, a perceived likelihood of violence

(PLV) score was calculated for each target by computing the mean rating of perceived violence across all participants (Stillman et al., 2010).

As previously stated, in exploring the evolutionary significance of a rapid threat-detection ability, Stillman et al. (2010) found that participants were able to perceive differences in level violence among the offenders and that both male and female participants were equally accurate in doing so. However, women rated targets as more violent overall. Moreover, they discovered a number of cues which participants relied on to inform their judgements. Some of these cues were valid indicators of violence, some were irrelevant, and some misled participants into making inaccurate decisions (Stillman et al., 2010). In sum, cues indicative of masculinity and high testosterone were related to accurate judgements, whereas facial emotion and good grooming were not good cues to more or less violence and often led to inaccurate judgements (Stillman et al., 2010).

Stillman et al.'s (2010) study was selected for replication for a number of reasons that make it a particularly strong test of the threat-detection abilities of humans. First, the measure of violence in this study is an actual history of engaging in violence, rather than an artificial measure of violence or aggression such as an in-lab task. Additionally, as targets' criminal records are publicly available, their past involvement in violence is certain and, therefore, there is no risk of socially-desirable responding to self-report measures of violence (Stillman et al., 2010). Another strength of this study design is the use of sex offenders as targets. Using this population allows for the categorization of targets into essentially matched groups that differ only in terms of their propensity for violence. That is, all target individuals engaged in risky, illegal, and socially unacceptable sexual behaviour that only differed in terms of associated violence, which limits the likelihood of explanations concerning differences in sexual deviance,

sex drive, or risk-taking (Stillman et al., 2010). Moreover, one argument that can be made is that all sex offenders are violent to some degree, regardless of the level of force or injury involved in their offences. As classifying any type of sex offender as ‘non-violent’ may be objectionable, offenders in the present study will be classified into ‘violent’ and ‘less-violent’ groups. Thus, the use of sexual offenders as targets in this study allows for a stringent test of humans’ abilities to detect what may be small differences in degree of violent propensity among individuals from a population of known violence. In addition, for the same reason, no purely non-violent group of individuals (e.g., non-criminals with no history of violence) is needed to test this hypothesis. This study was also selected for replication because of its use of static facial images as the central stimuli. Using static photographs allows for a specific analysis of the potential facial cues, including structural facial cues, that contribute to an individual’s perception of threat in others while eliminating many other potentially confounding variables (e.g., gestures, tone of voice, speech patterns, etc.). Moreover, this will also allow for a test of which of these facial cues, if any, are actually valid indicators of a violent propensity. The lack of available nuance provided by the presence of additional information also strengthens the test of humans’ threat detection capabilities as participants will be basing their judgements on a brief glance at a facial photograph.

For the current study, the ability of participants to distinguish between violent and less-violent sexual offenders from a two-second glance at photos of their faces was examined, as well as what facial cues participants attended to in order to inform this decision. Targets’ FWHR was measured alongside Stillman et al.’s (2010) list of cues as an additional non-verbal cue to masculinity. Appearance of mental disorder was also measured. Further extending the work of Stillman et al. (2010), the current study examined the relationship between rater personality

characteristics and accuracy in the detection of a propensity for violence. Rater characteristics that were examined included psychopathic personality traits, as conceptualized by two different self-report models (i.e., the SRP-4 and the TriPM), as well as the HEXACO dimensions of personality. This was examined across two samples, for which the methodologies of the study differed very slightly. This allowed for an examination of whether results differed between sample one, which presented the personality questionnaires in randomized order, and sample two, which did not. The research questions and hypotheses explored in the current study are outlined in the following section.

Research Question 1. Are individuals able to accurately distinguish between violent and less-violent sex offenders?

As it is believed that rapid threat-detection is a universal, evolutionarily-evolved mechanism that exists in all humans, it was expected that all participants would be able to accurately distinguish between the violent and less-violent sex offenders, consistent with the findings of Stillman et al. (2010).

Research Question 2: Are there gender differences in accuracy in distinguishing between violent and less-violent sex offenders? Do females perceive targets as more violent overall?

As previously stated, threat detection is believed to be a universal adaptation and, as such, males and females should be equally accurate. Consistent with the findings of Stillman et al. (2010), it was predicted that both males and females would perform equally well. This finding would also be consistent with the results of Chan et al. (2010), as accurately perceiving certain characteristics of a specific individual from a photograph of their face would constitute distinctive accuracy (i.e., no gender difference). It was also expected that women would, however, generally rate the targets as more violent, consistent with EMT (i.e., that less

detrimental mistakes may be made in the process of trying to avoid more detrimental mistakes; Haselton & Buss, 2000) and with Stillman and colleagues' (2010) findings.

Research Question 3: What facial cues do participants attend to in order to inform this decision (i.e., which cues are valid, poor, and misleading)?

The third research question concerned what valid, poor, and misleading facial cues, as rated separately by research assistants, participants attended to in order to inform their assessments of violence. It was expected that a number of cues indicative of level of testosterone would be used by participants to inform accurate assessments of inclination toward violence. Unfortunately, there is limited research on valid, poor, and misleading non-verbal cues to violence upon which to base specific predictions. However, as testosterone has been repeatedly linked to aggression and violence (e.g., Carré & McCormick, 2008, 2011; Salvador et al., 1999), facial cues that suggest a higher level of circulating testosterone should constitute valid cues to a propensity for violence (e.g., FWHR, facial masculinity, heavy brow). Valid cues, as operationalized by Stillman et al. (2010), are those cues that correlate with both a PLV by participants and a targets actual violent history. The specific testosterone-related valid cues to violence they discovered included facial masculinity, a heavier brow, apparent physical strength, and younger age. It was expected that these cues would, again, be valid cues signalling violence. Moreover, FWHR, as a measure of masculinity that has been linked to both levels of testosterone (Carré & McCormick, 2008; Carré et al., 2009; Carré et al., 2013; Short et al., 2012; Stillman et al., 2010; Stirrat & Perrett, 2010; Verdonck et al., 1999) and aggression (e.g., Carré & McCormick, 2008; Carré et al., 2009; Lefevre & Lewis, 2014), was also expected to be a valid cue to violence. Finally, considering the perception of the mentally-disordered as more likely to be violent (e.g., Field Institute, 1984), as well as the research evidence indicating that mentally-

disordered individuals are more likely to engage in violence (e.g., Arseneault et al., 2000; Monahan, 1992; Van Dorn et al., 2011), it was anticipated that appearance of a mental disorder would also be a valid cue to a propensity toward violence.

As defined by Stillman et al. (2010), poor cues are those that are not significantly related to either PLV or actual violent history. They found that attractiveness, smiling, facial hair, and the emotional expressions of surprise, fear, or sadness were all poor cues in their study (Stillman et al., 2010). With the exception of attractiveness, similar results were expected from the current study. However, as more attractive people tend to be evaluated more positively by others (e.g., Capetillo et al., 2008; Castellow et al., 1990; Curno, 2015; Esses & Webster, 1988; Langlois et al., 2000), it was anticipated that attractiveness would be a misleading cue in the current study, informing erroneous judgements of a lesser likelihood of violence. This result would also be consistent with DDT which posits that decision-making about the characteristics of others is influenced by lasting first impressions of potentially misleading cues (Porter & ten Brinke, 2009). In the context of this hypothesis, DDT suggests that participants will form a lasting impression from their two-second glance at an attractive person and, perhaps erroneously, judge that person to be less-violent.

According to Stillman et al. (2010), misleading cues are those which relate to PLV but not to actual violent history. In their study, they discovered that facial expressions of disgust and anger commonly influenced mistaken ratings of a higher likelihood for violence, while expressions of happiness, as well as good grooming, influenced mistaken ratings of lower likelihoods for violence (Stillman et al., 2010). Along with attractiveness, it was predicted that results from the current study would be consistent with these findings. Furthermore, the role of additional cues that were rated by research assistants for Stillman et al.'s (2010) study but were

not analyzed in the study results were evaluated. These include the presence of glasses, tattoos, or long hair, as well as facial asymmetry and creepiness. Moreover, ratings for appearance of age and appearance of mental disorder were collected. As research concerning non-verbal cues to violence is limited, the evaluation of the validity of these cues as indicators of violent propensity were exploratory.

Research Question 4: Are valid cues related to the ability to distinguish violent from less-violent targets?

It was hypothesized that the results of the current analyses, like those of Stillman and colleagues (2010), would suggest that valid cues were responsible for the ability to accurately identify targets with a history of violence.

Research Question 5: Are individuals with higher scores on measures of psychopathy more accurate in distinguishing between violent and less-violent sex offenders? Which of the facets of psychopathy, as measured by the SRP-4 and the TriPM, are most highly related to accuracy?

As individuals with higher levels of psychopathic traits have been demonstrated to be more adept in perceiving characteristics of others based on non-verbal cues (i.e., vulnerability; e.g., Book et al., 2013; Ritchie, 2016), participants who score higher on measures of psychopathic traits were expected to be better able to identify those sex offenders with a violent history from just a brief glance at digital images of their faces. In addition, previous research has found that all four facets of the SRP (i.e., IPM, CA, EL, and AB) are associated with an increased accuracy in judging violent victimization (Ritchie, 2016). However, the examination of how these facets, and the facets of the TriPM, relate to accuracy in the detection of violent propensity was exploratory.

Research Question 6: Which assessor personality traits (i.e., HEXACO) are associated with increased accuracy?

To date, no research has examined which HEXACO personality traits, if any, are associated with an increased ability to accurately detect a propensity for violence. As such, there were no specific hypotheses concerning potential relationships between the HEXACO personality dimensions and accuracy. Therefore, the analyses of the unique associations between the HEXACO personality dimensions and violence perception accuracy were exploratory.

Method

Participants and Sampling

Sample one. Male and female English-speaking undergraduate students from Carleton University were recruited for sample one through the SONA online participant recruitment bulletin (see Appendix A). Sample one consisted of 267³ students (189 females, 78 males) with a mean age of 19.8 (SD = 3.1). Of these participants, 73.6% were Caucasian (8.7% Asian, 5.7% Middle Eastern, 4.2% African-Canadian, 3.8% other, 1.9% Aboriginal/Native-Canadian/First Nations, 1.1% East Indian, 1.1% Hispanic/Latino). Those who participated were awarded course credit toward their first or second year psychology class.

Sample two. Using the same notice, male and female English-speaking undergraduate students from Carleton University were recruited through the SONA online participant recruitment bulletin. The total sample size for sample two was 203 (158 females, 44 males, 1 “other”) with a mean age of 20 (SD = 3.3). Of these participants, 64.5% were Caucasian (7.4% Asian, 6.4% Middle Eastern, 7.9% African-Canadian, 7.4% other, 2% Aboriginal/Native-Canadian/First Nations, 2.5% East Indian, 2% Hispanic/Latino). Students who participated were awarded course credit toward their first or second year psychology class.

Measures: Samples One and Two

Sex Offender Photographs. The facial photographs that were used in the current study are the same 84 images that were used by Stillman et al. (2010) in their investigation of whether or not individuals are able to accurately perceive a propensity for violence based solely on a two-second exposure to a digital image (Appendix B). As described previously, violent sex offenders

³ According to an a priori power analysis conducted in G*Power (Faul, Erdfelder, Lang, & Buchner, 2009), in order to achieve a projected power of 95% for a two-tailed correlational (Point-biserial) analysis with a medium effect size of .30 (Cohen, 1992), a sample size of 134 participants was required.

included those whose crimes involved an additional infliction of physical violence (i.e., rape with additional physical assault or weapon, sex offences accompanied by physical injury) and less-violent offenders included those whose crimes did not (i.e., non-violently touching a woman without permission, possessing child pornography, statutory offences). Fifty photographs of violent sex offenders and fifty photographs of less-violent sex offenders were collected from a public sex offender registry database in the United States, all of whom were of white males. Sixteen images were excluded due to either racial ambiguity, categorization difficulties, or poor image quality (Stillman et al., 2010).

Ratings of Non-Verbal Cues. Four separate raters, who were blind to the study hypotheses, provided separate ratings of the targets on a number of non-verbal cues (Appendix C). Two female research assistants and two male research assistants (two Caucasian, one Asian, and one Southeast Asian; ages ranging from 19 to 21) rated the photographs on the same cues, as well as some additional cues, that were used in Stillman et al.'s (2010) study. Prior to completing their individual ratings, a brief meeting with the raters took place to explain the cues they were expected to rate and to clarify any questions they had about the items. The majority of these cues were rated on Likert scales of 1 (*not at all*) to 9 (*very*). The others were rated on a yes or no basis (e.g., "Is the person in this photograph wearing glasses?").

Ratings from Stillman et al.'s (2010) design included three factors related to masculinity (i.e., heaviness of the brow, general facial masculinity, and apparent physical strength), attractiveness, whether the target is smiling, whether the target is wearing glasses, creepiness, facial asymmetry, the presence of facial hair, handicapped appearance, long hair, or tattoos, care in grooming, and the six basic facial expressions (i.e., anger, disgust, fear, happiness, sadness, and surprise; Ekman & Friesen, 1969). The additional ratings included the appearance of a

mental disorder, based on the research evidence presented above, as well as the appearance of arrogance, how old the target appeared to be, and the presence of jewellery, as exploratory variables. Ratings of the peripheral cue of clarity of the photo were also collected. These ratings were used to assess correlations between PLV by participants and actual violent histories of the offenders to determine which of these cues were valid (i.e., correlated with both PLV and violent history), poor (i.e., not correlated with either PLV or violent history), and misleading (i.e., correlated with PLV but not violent history).

Facial Width-to-Height Ratio. FWHR was calculated for each facial photograph by measuring the distance across the cheekbones (i.e., the distance between the left and right zygion, or the bizygomatic width) and dividing that value by the distance between the upper lip and the mid-brow (i.e., the upper facial height; e.g., Carré & McCormick, 2008; Carré et al., 2009; Carré et al., 2010; Carré et al., 2011; Short et al., 2012; Stirrat & Perrett, 2010). Two raters calculated the FWHR on all 84 photos to allow for an examination of interrater reliability.

Demographic Questionnaire. A short demographic questionnaire was used to collect information about the participants' ages, genders, and ethnicities (see Appendix D).

Attention Filters. For sample one, one quality control question (Appendix E) was randomly inserted into the questionnaires in order to help determine whether or not participants were paying adequate attention to the questions at hand and, therefore, whether they were responding appropriately. For sample two, three quality control questions were used (Appendix E). One of these attention filters was inserted randomly into each one of the personality questionnaires. These questions instructed participants to select a specific value when answering that question (e.g., "Select 2 when answering this question"). Participants who incorrectly answered all three of the attention filter questions were excluded from analyses.

HEXACO-60 (Ashton & Lee, 2009). The HEXACO-60 (see Appendix F) was constructed to be a quick and efficient method of measuring the six personality traits that are measured by the 100-item HEXACO Personality Inventory Revised: Honesty-Humility, Emotionality, Extraversion, Agreeableness versus Anger, Conscientiousness, and Openness to Experience (HEXACO-PI-R; Ashton & Lee, 2008; Lee & Ashton, 2004, 2006). Using results from a sample of college students and a sample of community adults, ten items were selected to represent each of the six scales by examining the primary and secondary factor loadings on each item in the HEXACO-Personality Inventory-Revised (HEXACO-PI-R). It was ensured that each of the four narrow facets of each scale were adequately represented within this selection (Ashton & Lee, 2009). The items take the form of a series of statements about the participant which he or she must rate, in terms of agreement, on a Likert-scale from 1 (*strongly disagree*) to 5 (*strongly agree*). Six scores are calculated by computing an average value for each subscale, representing the same six personality traits measured by the HEXACO-PI-R. The brevity of this instrument allows for researchers to measure these six major dimensions of personality under relative time restrictions (Ashton & Lee, 2009).

Lee and Ashton (2009) have demonstrated adequate convergent validity across student and community samples as the subscales of the HEXACO-60 produce moderate to high correlations when compared with their corresponding factors as measured by the NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1985). Additionally, internal consistencies remained high (i.e., α ranging from .73 to .80) and scale intercorrelations remained low ($<.30$) for this shortened measure of the HEXACO personality traits. Lee and Ashton (2009) concluded that the psychometric properties of the HEXACO-60 were highly consistent with those of the longer versions of the HEXACO-PI-R.

Self-Report Psychopathy Scale (SRP-4; Paulhus et al., 2016). The SRP-4 is a 64-item self-report measure of psychopathy that was derived from the PCL-R (Hare, 2003). Participants are presented with a variety of statements and are asked to respond with the extent to which they agree that each statement reflects their characteristics or actions. Their level of agreement is measured on a Likert-scale that ranges from 1 (*disagree strongly*) to 5 (*agree strongly*). Four subscale scores are calculated by summing 16 items each, resulting in scores for Interpersonal Manipulation (IPM), Callous Affect (CA), Erratic Lifestyle (ELS), and Criminal Tendencies (CT). The total score for this measure is calculated by summing these four subscales to get an overall score for psychopathy.

The Psychopathic Personality Inventory (PPI; Lilienfeld & Andrews, 1996) and the Levenson Self-Report Psychopathy Scale (LSRP; Levenson, Kiehl, & Fitzpatrick, 1995) have both been shown to correlate moderately with the SRP (Neal & Sellbom, 2012; Williams, Paulhus, & Hare, 2007), demonstrating reasonable convergent validity of the SRP with other self-report measures of psychopathy. Further, the SRP has been found to have high internal consistencies in both student and community samples (Mahmut, Menictas, Stevenson, & Homewood, 2011; Neal & Sellbom, 2012; Williams et al., 2007).

Brief Triarchic Psychopathy Measure (the Brief TriPM; Patrick, 2010). The Brief TriPM is a 58-item self-report questionnaire that was constructed to assess three personality constructs related to psychopathy: boldness, meanness, and disinhibition (Appendix G). The questionnaire presents personally-descriptive statements that participants are asked to rate with the response that best describes them on a Likert-scale of 1 (*true*) to 4 (*false*). Three separate scores can be computed for each of these domains by separately summing the 19 items designed to measure boldness, the 19 items designed to measure meanness, and the 20 items designed to

measure disinhibition (Patrick, 2010). These items were designed based on other recognized measures of psychopathy, including the PPI (Lilienfeld & Andrews, 1996; Lilienfeld & Widows, 2005) and the Externalizing Spectrum Inventory (ESI; Krueger, Markon, Patrick, Benning, & Kramer, 2007).

In an offender sample, scores on the three subscales of the TriPM have each been found to tap into separate facets of the PCL-R and, when taken together, to moderately relate to total scores on the PCL-R at a level comparable with the PPI ($>.5$; Patrick, 2010). Additionally, scores on the Brief TriPM were moderately associated with other self-report measures of psychopathy, namely the SRP-III, the PPI, and the Youth Psychopathic Traits Inventory (YPI; Andershed, Kerr, Stattin, & Levander, 2002) in a sample of college students (Patrick, 2010).

Accuracy Measure. Accuracy was measured by comparing raters' judgements of violence to targets' actual violent history. Accuracy scores were dichotomized, such that all ratings of 1 (*unlikely to be violent*) and 2 (*somewhat unlikely to be violent*) for less-violent targets were recoded as 1 (indicating accuracy) and all ratings of 3 (*somewhat likely to be violent*) and 4 (*likely to be violent*) were recoded as 0 (indicating inaccuracy). For the photographs of sex offenders who do have a violent history, ratings of 1 and 2 were recoded to have a value of 0 (an indicator of an inaccurate response) and ratings of 3 and 4 were recoded to have a value of 1 (an indicator of an accurate response). These values were summed in order to get an accuracy score that could range from 0 to 84.

Procedure

Prior to the collection of data, the current study was approved by the Carleton University Research Ethics Board-B. For each sample, undergraduate students from Carleton University were recruited through a sign-up posted on an online student participant pool. As the

convenience of an online study to gather data for a large sample size outweighs any potential disadvantages in this case (e.g., more subjective distractions) and an online modality does not increase socially-desirable responding, an online platform is a good alternative to in-laboratory administration (Clifford & Jerit, 2014). Participants who signed up were redirected to the study, which was delivered through Qualtrics online survey software. Upon redirection to the survey page on Qualtrics, participants were presented with an Informed Consent form, advising them of the purpose of the study and the tasks they would be completing and alerting them to any potential risks or benefits they may experience as a result of their participation in the study (Appendix H). As the purpose of this study was to evaluate individuals' capacities to detect small differences in propensity toward violence among a high risk group (i.e., sex offenders), participants were informed that the men in the photographs were all sex offenders. Those who did not consent were immediately redirected to the debriefing form (Appendix I).

Participants who consented to participate in the study were first presented with the demographic questionnaire, in which they were asked to provide information about their age, gender, and ethnicity. They were then instructed that they would be presented with the facial photographs of sex offenders which they would be asked to rate, on a scale of 1 (*unlikely to be violent*) to 4 (*likely to be violent*), in terms of propensity for violence (Appendix J). The 84 photographs were then presented to participants for two seconds each, in randomized order, and the participants were instructed to rate the individual in the photograph using the four-point Likert scale. After each rating, participants were provided with the option to withdraw their consent and cease participation. Participants who chose to withdraw were immediately redirected to the debriefing form and granted their partial course credit.

After all ratings were complete, participants in sample one were presented with the HEXACO-60 (Ashton & Lee, 2009), the SRP-III (Paulhus et al., 2016), and the Brief TriPM (Patrick, 2010), in this fixed order. In sample two, the same questionnaires were presented to participants upon completion of the photo ratings. However, they were presented in randomized order. Participants in both samples were instructed, prior to each questionnaire, that they had the option to skip any questions they did not wish to answer. On the bottom of each page, participants were provided with the option to withdraw their consent and cease participation with no penalty. Once the questionnaires were completed, participants were redirected to the debriefing form. The study took approximately 45 minutes to complete.

Results

Primary and Secondary Datasets

Due to the nature of the data, collection necessarily resulted in two datasets. The first, which will be referred to as the *primary dataset*, was organized by participants (i.e., the cases). This dataset included each participant's age, gender, ethnicity, ratings of target violence, and responses to the three self-report personality measures (i.e., HEXACO-60, SRP-4, and TriPM). Thus, participants' accuracy scores and total scores on the personality measures were in this dataset. As such, all analyses involving accuracy, gender, and/or personality traits were necessarily conducted in the primary dataset.

The *secondary dataset* was organized by target photograph (i.e., the cases). This dataset included the violence grouping (i.e., less-violent = 0, violent = 1), PLV score (calculated from the values in the primary dataset), FWHR value, and research assistants' cue ratings for each offender photograph. Therefore, all analyses involving the ability to distinguish between violent and less-violent targets (i.e., difference in PLV for violent and less-violent groups), PLV, and/or the facial cues were necessarily conducted using the secondary dataset.

Inter-Rater Reliability

Secondary dataset. As four raters were involved in the facial cue ratings, inter-rater reliability was conducted using Cronbach's Alpha (α), a measure of internal consistency (Stemler, 2004). Across all four raters, consistency ratings for most cues were considered acceptable ($\alpha > .70$; e.g., Tavakol & Dennick, 2011). Table 3 presents the reliability results across the individual ratings and across all the cues ($\alpha = .92$). Inter-rater agreement was acceptable for the majority of ratings. However, reliability coefficients for ratings of creepiness ($\alpha = .68$), facial masculinity ($\alpha = .62$), facial symmetry ($\alpha = .57$), and the emotional expressions

of disgust ($\alpha = .08$), fear ($\alpha = .65$), sadness ($\alpha = .55$), and surprise ($\alpha = .58$) were all below an acceptable value (i.e., $\alpha < .70$). The consistency values for most of these cues did not fall dramatically below the acceptable cutoff value and, as such, were included as they were critical to the analyses for the present study. However, as ratings of disgust expressions had very poor internal consistency ($\alpha = .08$), this cue was not examined in the analyses. Moreover, as none of the offenders had visible tattoos in their photographs, this rating was discarded.

FWHR. Two raters were involved in the measurements and calculations of each target's

Table 3

Internal Consistency Statistics for Inter-Rater Reliability Among Facial Raters

Rating	Reliability (α)
Smile	.91
Absence of glasses	1.00
Jewelry	.98
Facial hair	.95
Long hair	.93
Attractiveness	.86
Creepiness	.68
General facial masculinity	.62
Heaviness of brow	.71
Care in grooming	.73
Appearance of physical strength	.81
Quality of photo	.86
Appearance of physical handicap	.90
Appearance of mental disorder	.81
Facial symmetry	.57
Appearance of age	.94
Appearance of arrogance	.71
Angry expression	.78
Disgust expression	.08
Fear expression	.65
Happy expression	.92
Sad expression	.55
Surprise expression	.58
All variables	.92

Note. α = Cronbach's alpha.

FWHR. The raters' measurements were strongly positively correlated ($r = .81, p < .001$), indicating strong inter-rater agreement. As such, one set of measurements was chosen arbitrarily to represent this variable in analyses.

Data Preparation: Sample One

First, those participants who incorrectly responded to the attention check question were filtered out of the analyses. Of the 431 participants in the primary dataset, 267 participants correctly responded to the attention filter question. The reverse-scored items for each of the questionnaires were then reverse-coded and the data were examined for any out-of-range values. All values were within the appropriate range.

Missing data. The data were then analysed for missing values. Five items were missing from the HEXACO-60. Each of these items corresponded to a different dimension. Thus, to calculate accurate total scores for these dimensions, a mean was calculated using the available dimension items if responses to at least 80% of the items were available. Of the variables, 12 were missing data. The amount of missing data on these variables ranged from 0% to 18%. The percentage of missing item data for each participant ranged from 0% to 69.7% (0 to 124 items) and missing variable data ranged from 0% to 73.3 % (0 to 11 variables). The missing data pattern was examined more formally by using Little's MCAR test (Little, 1988). This value was non-significant ($p < .05$), suggesting that the data may be missing at random.

Next, for each questionnaire, participants with missing data were compared to those without. This was done by first calculating a missing values score for each participant for each questionnaire. These scores were then dichotomized to represent those with missing values and those without for each of the HEXACO-60, SRP-4, and TriPM. Using t-tests and chi squares, these groups were compared on both age and gender to determine if they differed in terms of

either of these variables. None of these tests produced significant results. The difference between the ages of participants with missing data on the HEXACO-60 and those without approached significance ($p = .05$) but was not below a minimum p -value of .05 and, thus, was not statistically significant. Based on these results, the data can be considered missing at random.

As the analyses suggested that the data were missing at random, multiple imputation was used to estimate the missing values. Due to the fact that one cannot be certain what the correct imputed value should be for missing values, the multiple imputation procedure (Rubin, 1987) produces a random sample of many plausible values for the missing data, resulting in multiple datasets (m datasets). This accounts for any potential error in imputed values (Yuan, 2010). All primary analyses are then performed on the m complete datasets and their results are combined for final, valid inferences to be made. For the current sample, 50 iterations ($m = 50$) were used to create the imputed datasets, resulting in 50 complete datasets.

Linearity and homoscedasticity. The linearity and homoscedasticity among the variables were then examined. This was done by examining bivariate scatterplots of each combination of variable pairs. All combinations of variables satisfied these conditions.

Univariate outliers and normality. The data were then examined to identify the presence of any univariate outliers. Upon visual inspection of boxplots and histograms, it appeared that there may be outliers present on a number of variables. To inspect this more thoroughly, z -scores were calculated for all variables. Those z -score values which exceeded a value of ± 3.29 were considered to be extreme outliers (Tabachnick & Fidell, 2013). Using this procedure, 6 extreme outliers were identified. HEXACO agreeableness, SRP-4 Interpersonal Manipulation, TriPM Disinhibition, and TriPM Total each had one extreme value. SRP-4 Criminal Tendencies had two extreme outliers.

The data were then screened for departures from normality. Skewness and kurtosis become an issue when their z values exceed ± 1.96 (i.e., statistic $\div$ standard error $> \pm 1.96$; Tabachnick & Fidell, 2013). According to this guideline, Emotionality, Agreeableness, and Conscientiousness scores on the HEXACO-60, and Meanness, Disinhibition, and Total scores on the TriPM were significantly negatively skewed⁴. Interpersonal Manipulation, Callous Affect, Criminal Tendencies, and Total scores on the SRP-4 were significantly positively skewed⁵. Significant kurtosis was only an issue for Criminal Tendencies on the SRP-4.

In attempt to resolve the issues of the presence of extreme outliers and skewness, logarithm transformations were conducted. These transformations fixed the outlier problem for most of the variables except for one of the outliers associated with SRP-4 Criminal Tendencies. This case was removed from the dataset and all z -scores were examined again to ensure that all outliers had been eliminated. All outliers had been resolved. The transformations resolved the skewness issues for all variables except for Interpersonal Manipulation and Criminal Tendencies on the SRP-4. Inverse transformations were attempted for these two variables but these also had no impact on the skew. Thus, the original, moderate, transformations were used in the primary analyses for these two variables, as were the transformed data for the other skewed variables. Moreover, the sample size was large enough that the data were robust to violations of normality (i.e., $n = 267 > 40$; Field, 2013).

Multicollinearity. Multicollinearity was examined by evaluating the values of the variance inflation factor (VIF) and tolerance for each variable (Field, 2013). VIFs indicated that

⁴ *Note.* Responses to the TriPM items are on a scale of 1 (*True*) to 4 (*False*). Thus, a greater amount of higher scores indicates lower levels of psychopathy. This is to be expected in a sample of undergraduate university students.

⁵ *Note.* Responses to the SRP-4 are on a scale of 1 (*Disagree Strongly*) to 5 (*Agree Strongly*). Thus, a greater amount of lower scores indicates lower levels of psychopathy. This is to be expected in a sample of undergraduate university students.

multicollinearity was not a concern as none of these values exceeded 10. Moreover, all tolerance values were greater than 0.20, indicating that multicollinearity was not an issue (Field, 2013).

Descriptive statistics and internal consistency. Internal consistencies for all of the dimensions on the self-report questionnaires were acceptable. Descriptive statistics and internal consistency values (Cronbach's Alpha) for each dimension of each measure are presented in Table 4. Bivariate correlations between the HEXACO-60 dimensions and the dimensions of the psychopathy measures are displayed in Table 5 and bivariate correlations between the dimensions of the two psychopathy measures (i.e., SRP-4 and TriPM) are displayed in Table 6.

Table 4

Sample One: Descriptive Statistics and Internal Consistencies for the HEXACO-60, SRP-4, and TriPM

Measure	Dimension/Factor/Facet	Mean	Standard Deviation	Internal Consistency (α)
HEXACO-60	Honesty-Humility	3.20	0.68	.77
	Emotionality	3.34	0.68	.79
	Extraversion	3.21	0.64	.75
	Agreeableness	3.13	0.61	.79
	Conscientiousness	3.56	0.58	.76
	Openness	3.36	0.66	.75
SRP-4	Interpersonal Manipulation	39.04	9.85	.85
	Callous Affect	37.10	8.83	.81
	Erratic Lifestyle	42.32	10.23	.84
	Criminal Tendencies	26.07	7.09	.73
	SRP-4 Total	144.41	28.76	.92
TriPM	Boldness	47.13	8.62	.83
	Meanness	63.36	9.06	.90
	Disinhibition	63.26	9.17	.81
	TriPM Total	58.10	6.51	.91

Note. SRP-4 = Self-Report Psychopathy Scale (Paulhus et al., 2016); TriPM = Brief Triarchic Psychopathy Measure (Patrick, 2010); α = Cronbach's alpha.

Table 5

Sample One: Bivariate Correlations of the SRP-4 and the TriPM with the HEXACO-60

Psychopathy Measure	H	E	X	A	C	O
SRP-4						
Interpersonal Manipulation	-.55***	.38***	.00	.38***	.19	-.02
Callous Affect	-.43***	.58***	-.03	.41***	.19	-.21***
Erratic Lifestyle	-.37***	.35***	.06	.23***	.44***	-.02
Criminal Tendencies	-.40***	.19*	-.01	.22***	.17**	.05
SRP-4 Total	-.52***	.49***	.03	.36***	.35***	-.06
TriPM						
Boldness	.14*	-.66***	-.60***	.03	.18**	-.07
Meanness	-.52***	.47***	-.07	.44***	.29***	-.20**
Disinhibition	-.41***	.04	-.18**	.28***	.51***	-.01
TriPM Total	-.48***	.53***	.16*	.32***	.25***	-.01

Note. H = Honest-humility E = Emotionality; X = Extraversion; A = Agreeableness; C = Conscientiousness; O = Openness; SRP-4 = Self-Report Psychopathy Scale (Paulhus et al., 2016); TriPM = Triarchic Psychopathy Measure (Patrick, 2010).

* $p < .05$. ** $p < .01$. *** $p \leq .001$

Table 6

Sample One: Bivariate Correlations between the SRP-4 with the TriPM

	TriPM			
	Boldness	Meanness	Disinhibition	TriPM Total
SRP-4				
Interpersonal Manipulation	-.40***	-.67 ***	-.46***	-.67***
Callous Affect	-.39***	-.78***	-.40***	-.71***
Erratic Lifestyle	-.33***	-.57***	-.66***	-.70***
Criminal Tendencies	-.18**	-.45***	-.54***	-.55***
SRP Total	-.42***	-.78***	-.65***	-.83***

Note. SRP = Self-Report Psychopathy Scale (Paulhus et al., 2016); TriPM = Triarchic Psychopathy Measure (Patrick, 2010).

** $p < .01$. *** $p < .001$.⁶

⁶ Higher facet scores on the TriPM indicate lower levels of those personality traits, whereas higher facet scores on the SRP indicate higher levels of those personality traits. This explains the negative direction of the correlations between the two measures.

Primary Analyses: Sample One

Distinguishing target violence. A Perceived Likelihood of Violence (PLV) score was first computed for all targets, calculated as a mean score across all respondents' ratings of violent propensity and saved in the secondary dataset. This score was used to conduct preliminary analyses to examine whether participants rated violent offenders as more violent than the less-violent offenders (i.e., a comparison between the mean PLV for violent and the mean PLV for less-violent offenders). To determine whether participants were able to distinguish between violent and less-violent sexual offenders based solely on their facial characteristics, a t-test was conducted in the secondary dataset. It was hypothesized that participants would be able to distinguish between violent and less violent offenders (e.g., Bar et al., 2006; Carré et al., 2009; Stillman et al., 2010; Willis & Todorov, 2006). The mean PLV scores for the violent offenders and the less-violent offenders were compared to determine whether participants rated violent offenders higher on violent propensity than less-violent offenders. The results indicated that there was no significant difference between the PLV scores for the violent ($M = 2.66$, $SD = 0.32$) and less-violent ($M = 2.62$, $SD = 0.37$) offenders ($t(82) = -0.57$, $p = .57$).

Gender differences in accuracy. To determine whether one gender was more accurate in their assessments of violence than the other, and whether females perceived targets as more violent overall, a one-way ANOVA was conducted in the primary dataset. An average PLV score was first calculated for each participant across all target photos. Sex differences in accuracy scores, as well as in participants' average PLV scores, were then examined. It was expected that males and females would perform equally well in terms of accuracy as threat detection is purported to be a universal adaptation (e.g., McArthur & Baron, 1983; Oosterhof & Todorov, 2008; Sell et al., 2008; Short et al., 2012). Moreover, this finding would be consistent with the

results of Stillman et al. (2010) and Chan et al. (2010), as described previously. Additionally, it was hypothesized that, consistent with EMT (Haselton & Buss, 2000) and Stillman et al. (2010), females would rate the targets as more violent overall (i.e., across violent and less-violent groups) than males.

The results of the ANOVA were consistent with previous research and supported the hypothesis that both sexes would be equally accurate in their assessments of violence, as there was no significant effect of gender on accuracy, $F(1, 265) = 2.64, p = .11$. However, the hypothesis that females would rate targets as more violent across both groups of offenders was not supported. That is, there was no significant effect of gender on overall average PLV ratings, $F(1, 265) = 0.83, p = .36$.

Valid, poor, and misleading facial cues. First, to have only one facial rating per cue for each offender (e.g., facial masculinity), a mean was calculated across all four raters for each cue. Based on previous research (Stillman et al., 2010; Stillman & Maner, 2009), potential cues were initially sorted into three groups based on bivariate correlations between the cues and both PLV and actual violent history: valid, poor, or misleading. Valid cues were those that were correlated with both PLV and actual violent history, poor cues were those that were not significantly correlated with either PLV or actual violent history, and misleading cues were those that correlated with PLV but not actual violent history. However, during analyses, a fourth group was discovered in which the cues did not correlate with PLV but did correlate with an actual history of violence. The resulting grouping designations were as follows: valid (detected; i.e., correlated with actual violent history and PLV), valid (undetected; i.e., correlated with actual violent history but not PLV), poor, and misleading.

To examine which cues were valid (detected), valid (undetected), poor, and misleading in participants' judgements of violence, bivariate correlations were conducted between each of the cues and PLV. It was hypothesized that cues suggestive of higher levels of testosterone, including facial masculinity, heavy brow, apparent physical strength, younger age (e.g., Carré & McCormick, 2008, 2011; Salvador et al., 1999; Stillman et al., 2010), and FWHR (e.g., Carré & McCormick, 2008; Carré et al., 2009; Lefevre & Lewis, 2014; Short et al., 2012; Stirrat & Perrett, 2010), as well as the appearance of mental disorder would be valid (detected) cues, significantly correlating with both PLV and actual history of violence (e.g., Arseneault et al., 2000; Monahan, 1992; Van Dorn et al., 2011). Poor cues are those facial characteristics that were not significantly correlated with either PLV or an actual history of violence and were hypothesized to include smiling, facial hair, and the emotional expressions of surprise, fear, and sadness (Stillman et al., 2010). Misleading cues were those which correlated significantly with PLV but not with an actual history of violence. It was predicted that, in the current sample, misleading cues would include good grooming and attractiveness as well as the facial expressions of disgust, anger, and happiness. However, as ratings of disgust expressions had such poor reliability, this cue was not examined. No other specific hypotheses were made about the additional cues measured (i.e., absence of glasses, long hair, facial asymmetry, and creepiness). Thus, analyses of the relationships between each of these cues and both PLV as well as actual violence were exploratory. These correlations are presented in Table 7. For correlations between cues, actual violence, and participants' PLV ratings by gender for sample one, see Appendix K.

As previously noted, this study is only the second to examine the relationships between very brief exposures to static facial cues and resulting judgements of violence and accuracy of

Table 7

Sample One: Bivariate Correlations of Facial Cues with PLV and Actual Violent History

Facial Cue	PLV	Violent History
FWRH	.11	-.14
Smile	.12	-.13
Absence of glasses	.22*	.22*
Jewelry	-.03	-.14
Facial hair	-.12	-.03
Long hair	.11	.10
Attractiveness	.10	.37***
Creepiness	.21	-.27*
General facial masculinity	.24*	.28*
Heaviness of brow	.35***	.31**
Care in grooming	-.10	.38***
Appearance of physical strength	.36***	.11
Quality of photo	-.27*	-.15
Appearance of physical handicap	-.17	-.25*
Appearance of mental disorder	-.02	-.32**
Facial symmetry	-.20	.31**
Older age	-.43***	-.23*
Appearance of age	-.35***	-.26*
Appearance of arrogance	.26*	.10
Angry expression	.41***	-.12
Fear expression	-.15	.11
Happy expression	-.13	.03
Sad expression	-.02	.15
Surprise expression	-.19	-.05

Note. PLV = Perceived likelihood of violence.

* $p < .05$. ** $p \leq .01$. *** $p \leq .001$.

threat detection. Thus, potential family-wise error was not corrected for and a cutoff value of $p = .05$ was used to evaluate the statistical significance of results. As expected, correlational analyses from the current sample identified general facial masculinity (i.e., more violence), heaviness of brow (i.e., more violence), and older age (i.e., less violence) as valid (detected) cues to level of violence. Of the exploratory variables, the appearance of older age, which was, not surprisingly, strongly correlated with actual older age ($r = .91, p < .01$), was also a valid (detected) cue (less violence). Interestingly, the absence of glasses also correlated with both PLV and actual history

of violence, suggesting that not having glasses was a valid (detected) cue to violence in this sample.

Although they were predicted to be misleading cues to a lower propensity toward violence, the results from sample one suggest that attractiveness and good grooming were both valid (undetected) cues toward violent propensity. These cues significantly correlated with each other ($r = .60, p < .01$), and they both correlated with an actual history of violence, but neither was correlated with PLV. Of the exploratory variables, facial symmetry was also a valid (undetected) cue, as participants did not rely on this cue in their assessments of violence even though it was positively correlated with a history of violence. Further, contrary to prediction, the appearance of a mental disorder was not found to be a valid (detected) cue. Rather, it was found to be a valid (undetected) cue, as it was not correlated with perceptions about target violence but it was correlated with a less violent history. This was also the case for the appearance of a physical handicap. Interestingly, creepiness was negatively correlated with violence, suggesting that creepier looking targets were actually less violent. However, creepiness was not correlated with PLV in sample one, indicating that creepiness was a valid (undetected) cue to less actual violence.

As predicted, smiling, facial hair, and the emotional expressions of surprise, fear, and sadness were all found to be poor cues, correlating with neither PLV nor violent history. The presence of jewelry and the presence of long hair, which were both exploratory variables, were also poor cues. Moreover, FWHR was found to be a poor cue, contrary to the initial hypothesis that this metric would be a valid indicator of actual violent history. The expression of happiness was also a poor cue, contrary to the prediction that it would be a misleading cue to ratings of less violent propensity.

Consistent with expectations, the expression of anger was found to be a misleading cue. Similarly, of the exploratory variables, the appearance of arrogance was found to be a misleading cue as it was correlated with perceptions of a higher propensity for violence but was not correlated with an actual history of more violence. These results suggest that individuals relied on angry expressions and arrogant appearances to erroneously inform ratings of higher propensities toward violence. Interestingly, photo quality, while not a facial cue, was also found to be a misleading cue, correlating with level of perceived violence but not with an actual history of violence. Specifically, targets in lower quality photos were rated as being less violent than targets in higher quality photos. Moreover, the appearance of physical strength, although it was hypothesized to be a valid cue to more violence, was found to be a misleading cue, correlating with higher PLVs but not with actual history of violence. Valid (detected), valid (undetected), poor, and misleading cues from sample one are summarized in Table 8.

Table 8

Sample One: Valid, Poor, and Misleading Facial Cues in Sample One

Valid (Detected)	Valid (Undetected)	Poor	Misleading
Older age*	Facial symmetry	FWHR	Angry expression
Appearance of age*	Appearance of mental	Smile	Appearance of
Absence of glasses	disorder	Jewelry	physical strength
General facial	Appearance of	Facial hair	Appearance of
masculinity	physical handicap	Long hair	arrogance
Heaviness of brow	Care in grooming	Happy expression	Quality of photo
	Attractiveness	Surprise expression	
	Creepiness	Fear expression	
		Sad expression	

Note. * = cues that were correlated with less propensity for violence. FWHR = facial width-to-height ratio.

Valid cues and decision making. As actual violent history was dichotomized, a hierarchical binary logistic regression was used to examine the contribution of valid (detected) cues to accuracy in distinguishing between violent and less-violent targets. It was hypothesized that valid cues (i.e., valid detected cues) would contribute to participants' accuracy scores and that poor and misleading cues would contribute little to this ability (Stillman et al., 2010). Actual violent history was the dependent variable in these analyses. In step one of this regression, PLV was entered as the independent variable. The results from this step demonstrated that PLV was not a significant predictor of actual history of violence (Wald = 0.33, $p = .57$). In step two of this regression, the cues identified in the previous analyses as valid (detected; i.e., older age, appearance of older age, absence of glasses, general facial masculinity, and heavy brow) were entered. Entering these cues did not cause a statistically significant change in the relationship between PLV and actual violent history (Wald = 2.06, $p = .15$). However, it should be noted that heaviness of brow (Wald = 4.01, $p = .05$) and facial masculinity (Wald = 3.46, $p = .06$) both approached significance.

Three similar regressions were conducted to examine whether valid (undetected) cues, poor cues, and misleading cues contributed to the ability to distinguish between violent and less-violent targets. Similar results were found from these analyses with both valid (undetected) and poor cues. However, creepiness (Wald = 6.12, $p = .01$), although it was not used by participants to inform their judgements of violence, was significantly related to a history of violence.

Taken together, the results suggest that participants were not able to distinguish between violent and less-violent targets. None of the facial cues tested in sample one contributed to such an ability. However, the results suggest that a few of these cues may be predictors of actual

violence, without contributing to an ability to actually decipher between more and less violent individuals (e.g., creepiness).

Psychopathy and accuracy. Previous research has demonstrated an enhanced ability in individuals with higher levels of psychopathic traits, compared to those with lower levels of psychopathic traits, to more accurately identify certain characteristics of others based on non-verbal cues (e.g., Book et al., 2013; Ritchie, 2016). As such, based on the assumption that these individuals are generally better able to perceive the characteristics of others, it was predicted that they would also be more adept at judging violent propensity of the targets in the current study based on their facial features. Examination of which of the conceptual facets of psychopathy, as conceptualized by two different structures (i.e., the four factor model of the SRP-4 and three factor model of the TriPM), contributed most to accuracy in these perceptions was exploratory as this has not previously been examined with respect to the detection of violence.

Bivariate correlations were first conducted between the facet-level and total scores on measures of psychopathy and the participant accuracy scores (i.e., how often the participants were accurate in violent assessments, ranging from 0-84) to determine if level of psychopathy was related to accuracy (see Table 9). Two measures that are based on differing conceptions of the structure of psychopathy (i.e., the SRP-4 is based on a 4 factor conceptualization and the TriPM is based on a 3 factor conceptualization) were used to allow for an examination of which facets in each factor structure, if any, are related to accuracy in threat detection. Results from these preliminary analyses demonstrated that none of the facets or total scores on measures of psychopathy were correlated with accuracy in distinguishing between violent and less-violent offenders.

Given these findings, it is not surprising that the results of the multiple regression

Table 9

Sample One: Bivariate Correlations between Each SRP-4 and TriPM Personality Dimension and Accuracy

Psychopathy Measure	Accuracy
SRP-4	
Interpersonal Manipulation	-.10
Callous Affect	.02
Criminal Tendencies	-.03
Erratic Lifestyle	-.02
SRP Total	-.04
TriPM	
Boldness	.04
Meanness	.00
Disinhibition	-.01
TriPM Total	.01

Note. SRP = Self-Report Psychopathy Scale (Paulhus et al., 2016); TriPM = Triarchic Psychopathy Measure (Patrick, 2010).

including the facet and total scores as independent variables suggested that this was not a good model ($R^2 = 0.05$, $F(9, 177) = 0.99$, $p = .45$). None of the facet or total scores for the psychopathy measures significantly predicted accuracy, with the exception of the callous affect facet of the SRP ($\beta = 0.83$; $t(9) = 2.07$, $p = .04$). Separate models including (1) only the TriPM facet and total scores ($R^2 = 0.01$, $F(4, 214) = 0.27$, $p = .90$) and (2) only the SRP facet and total scores ($R^2 = 0.02$, $F(5, 211) = 0.80$, $p = .55$) as independent variables also were not good models for the prediction of participant accuracy scores. In these regressions, none of the individual facet scores were significant predictors of accuracy.

General personality and accuracy. While there is existing research evidence to suggest that personality influences individuals' person perception abilities (e.g., Ritchie, 2014, 2016), research has yet to examine the unique roles of the HEXACO personality traits in the accuracy of rapid judgements of threat based on non-verbal cues. As such, no specific a priori hypotheses were made concerning these traits and threat-detection accuracy. These relationships were first

explored by conducting bivariate correlations between each of the HEXACO personality facet scores and participants' total accuracy scores (i.e., how often the participants were accurate in violent assessments, ranging from 0-84; see Table 10).

Results from these analyses demonstrated that none of these facets were related to participants' accuracy in distinguishing between violent and less-violent offenders. Thus, it is not surprising that a multiple regression analysis including all six of these facets as predictor variables for accuracy produced results that were statistically insignificant [$R^2 = 0.01$, $F(6, 255) = 0.28$, $p = .95$] and that none of the facet scores were demonstrated to individually significantly predict accuracy.

Table 10

Sample One: Bivariate Correlations between Each HEXACO-60 Personality Dimension and Accuracy

HEXACO-60 Dimension	Accuracy
Honesty-Humility	-.01
Emotionality	.02
eXtraversion	.06
Agreeableness	-.01
Conscientiousness	-.03
Openness	.05

Data Preparation: Sample Two

Participants who incorrectly responded to all three of the attention check questions in sample two were filtered out of the analyses. Of the 287 participants in the primary dataset, 203 participants correctly answered these questions. Next, all of the items requiring reverse-scoring for each personality questionnaire were reverse-coded and the data were examined for potential out-of-range values. All values were within the appropriate range. As in sample one, the total scores for the HEXACO-60 were calculated by computing a mean score using the available dimension items, only if responses to at least 80% of the items were available.

Missing data. The data for sample two were then analysed for missing values. All of the variables were missing some data, ranging from 4.4% to 23.6%. The percentage of missing item data for each participant ranged from 2.5% to 9.4% (5 to 19 items) and missing variable data ranged from 6.7% to 100% (1 to 15 variables). The missing data pattern was then examined more formally using Little's MCAR test (Little, 1988). The statistic from the Little's MCAR test was non-significant ($p < .05$), suggesting that the data may be missing at random.

Participants with missing data on each questionnaire were then compared to those without. As in sample one, a missing values score was calculated and dichotomized to represent those with missing values and those without for each participant for each questionnaire. To determine if these groups differed, they were compared on both age and gender using t-tests and chi squares. None of these tests produced significant results. Age differences between participants with missing data on the HEXACO-60 and those without approached significance ($p = .05$) but were not below a p value of .05. These results suggest that the missing data in sample two were missing at random. Thus, as the analyses suggested that the data were missing at random, multiple imputation was used to estimate the missing values, as was done for sample

one. All primary analyses were performed on the imputed dataset for accurate results. For sample two, 50 iterations ($m = 50$) were used to produce 50 complete datasets.

Linearity and homoscedasticity. The linearity and homoscedasticity among the variables were then examined by examining bivariate scatterplots of each combination of variable pairs. All variable pairs satisfied these conditions.

Univariate outliers and normality. The data were then examined for univariate outliers. Visually examining boxplots and histograms, it appeared that there may be outliers present on a few variables. Z-scores were computed for all variables to investigate this more thoroughly. Those z-score values which exceeded a value of ± 3.29 were considered to be extreme outliers (Tabachnick & Fidell, 2013). Four univariate outliers were identified using this procedure. HEXACO-60 extraversion and conscientiousness each had one extreme value, and SRP-4 Criminal Tendencies had two extreme outliers.

The normality of the data was then examined. As previously noted, skewness and kurtosis present concerns when their z values exceed ± 1.96 (i.e., statistic $\div$ standard error $> \pm 1.96$; Tabachnick & Fidell, 2013). According to this guideline, HEXACO-60 Agreeableness and TriPM Meanness, Disinhibition, and Total scores were all significantly negatively skewed⁷. Callous Affect, Criminal Tendencies, and Total scores on the SRP-4 were significantly positively skewed⁸. Significant kurtosis was not an issue.

Log transformations were conducted in order to resolve the issues of extreme outliers and significant skews. These rectified the skewness problems for all of the negative skews and for all

⁷ Note. Responses to the TriPM items are on a scale of 1 (*True*) to 4 (*False*). Thus, a greater amount of higher scores indicates lower levels of psychopathy. This is to be expected in a sample of undergraduate university students.

⁸ Note. Responses to the SRP-4 are on a scale of 1 (*Disagree Strongly*) to 5 (*Agree Strongly*). Thus, a greater amount of lower scores indicates lower levels of psychopathy. This is to be expected in a sample of undergraduate university students.

of the positive skews except for SRP-4 criminal tendencies. An inverse transformation was attempted for this variable but did not fix the problem. Thus, the original moderate transformation was used for the primary analyses, as were the transformed data for the other skewed variables. Moreover, the sample size was large enough that the data were robust to violations of normality (i.e., $n = 203 > 40$; Field, 2013).

Multicollinearity. To examine multicollinearity, the values of the variance inflation factors (VIF) and tolerance for each variable were examined (Field, 2013). The VIFs indicated that multicollinearity was not a concern for any of the variables (i.e., $VIF < 10$, tolerance $> .20$).

Descriptive statistics and internal consistency. Internal consistencies for the dimensions on each of the self-report questionnaires were acceptable. Descriptive statistics and internal consistency values (Cronbach's Alpha) for each dimension of each measure are summarized in Table 11. Bivariate correlations between the HEXACO-60 dimensions and the dimensions of the psychopathy measures are displayed in Table 12 and bivariate correlations between the dimensions of the two psychopathy measures are displayed in Table 13.

Table 11

Sample Two: Descriptive Statistics and Internal Consistencies for the HEXACO-60, SRP-4, and TriPM

Measure	Dimension/Factor/Facet	Mean	Standard Deviation	Internal Consistency (α)
HEXACO-60	Honesty-Humility	3.18	0.65	.74
	Emotionality	3.40	0.66	.79
	Extraversion	3.13	0.66	.78
	Agreeableness	3.10	0.58	.73
	Conscientiousness	3.63	0.61	.83
	Openness	3.32	0.65	.76
SRP-4	Interpersonal Manipulation	39.63	10.26	.85
	Callous Affect	36.59	9.71	.82
	Erratic Lifestyle	41.36	11.00	.83
	Criminal Tendencies	27.98	8.57	.80
	SRP-4 Total	144.28	31.84	.94
TriPM	Boldness	47.67	7.95	.74
	Meanness	62.50	9.37	.86
	Disinhibition	63.37	9.57	.88
	TriPM Total	57.95	6.71	.91

Note. SRP-4 = Self-Report Psychopathy Measure (Paulhus et al., 2016); TriPM = Brief Triarchic Psychopathy Measure (Patrick, 2010); α = Cronbach's Alpha.

Table 12

Sample Two: Bivariate Correlations of the SRP-4 and the TriPM with the HEXACO-60

Psychopathy Measure	H	E	X	A	C	O
SRP-4						
Interpersonal Manipulation	-.64***	-.38***	-.11	.41***	-.49***	-.05
Callous Affect	-.42***	-.55***	-.15	.32***	-.29***	.00
Erratic Lifestyle	-.54***	-.41***	-.03	.24***	-.59***	.08
Criminal Tendencies	-.26***	-.24**	-.04	.04	-.45***	.00
SRP-4 Total	-.62***	-.50***	-.08	.35***	-.55***	.03
TriPM						
Boldness	.02	.63***	-.52***	.11	.03	-.06
Meanness	-.52***	-.42***	-.12	.38***	-.54***	-.10
Disinhibition	-.52***	-.14	-.13	.35***	-.69***	.07
TriPM Total	-.48***	-.54***	.13	.31***	-.57***	.04

Note. H = Honest-humility; E = Emotionality; X = Extraversion; A = Agreeableness; C = Conscientiousness; O = Openness; IPM = Interpersonal manipulation; CA = Callous affect; ELS = Erratic lifestyle; CT = Criminal tendencies; SRP-4 = Self-Report Psychopathy Scale (Paulhus et al., 2016); TriPM = Triarchic Psychopathy Measure (Patrick, 2010).

** $p < .01$. *** $p < .001$.

Table 13

Sample Two: Bivariate Correlations between the SRP-4 and the TriPM

	TriPM			
	Boldness	Meanness	Disinhibition	TriPM Total
SRP-4				
Interpersonal Manipulation	-.29***	-.70***	-.63***	-.75***
Callous Affect	-.30***	-.70***	-.42***	-.67***
Erratic Lifestyle	-.28***	-.57***	-.71***	-.73***
Criminal Tendencies	-.13	-.50***	-.54***	-.53***
SRP Total	-.31***	-.78***	-.74***	-.86***

Note. IPM = Interpersonal manipulation; CA = Callous affect; ELS = Erratic lifestyle; CT = Criminal tendencies; SRP = Self-Report Psychopathy Scale (Paulhus et al., 2016); TriPM = Triarchic Psychopathy Measure (Patrick, 2010).

*** $p < .001$.⁹

⁹ Higher facet scores on the TriPM indicate lower levels of those personality traits, whereas higher facet scores on the SRP indicate higher levels of those personality traits. This explains the negative direction of the correlations between the two measures.

Primary Analyses: Sample Two

Distinguishing target violence. Using the same method as in sample one, a PLV score was first computed for all targets. Using this score, a preliminary analysis exploring whether violent offenders were rated as more violent than less-violent offenders was conducted. Specifically, a t-test was conducted in the secondary dataset to determine whether or not participants were able to distinguish between violent and less-violent offenders based on their overt facial characteristics (i.e., a comparison of the mean PLV for violent and the mean PLV for less-violent offenders). Again, it was hypothesized that participants would be able to distinguish between the groups of offenders (e.g., Stillman et al., 2010). Consistent with results from sample one, the results of this test demonstrated that there was no significant difference between the PLV scores for the violent ($M = 2.64$, $SD = 0.31$) and less-violent ($M = 2.60$, $SD = 0.36$) offenders ($t(82) = -0.52$, $p = .60$).

Gender differences in accuracy. As with sample one, to determine whether one gender was more accurate in their assessments of violence than the other and whether females perceived targets as more violent overall, a one-way ANOVA was conducted in the primary dataset for sample two. An average PLV score was calculated for each participant across all target photos and sex differences in accuracy scores and in the average PLV scores were then tested. Again, it was hypothesized, based on the assumption that threat detection is a universal adaptation in humans (e.g., McArthur & Baron, 1983; Oosterhof & Todorov, 2008; Sell et al., 2008; Short et al., 2012), that both genders would be equally accurate in the detection of violence. This result would also be consistent with the results of Stillman et al. (2010) and Chan et al. (2010). It was also expected that females would rate the targets as more violent across both groups of offenders

(i.e., violent and less-violent groups) than males, consistent with EMT (Haselton & Buss, 2000) and the results of Stillman et al. (2010).

The results of the ANOVA from sample two replicated the results from sample one. Moreover, they were consistent with previous research and supported the hypothesis that males and females would not differ in level of accuracy with respect to judgements of violence, as there was no significant effect of gender on accuracy ($F [2, 200] = 0.58, p = .56$). Again, the results of the ANOVA did not support the hypothesis that females would rate targets as more violent across violent and less-violent offenders as there was no significant effect of gender on average PLV ratings ($F [2, 200] = 0.35, p = .71$).

Valid, poor, and misleading facial cues. As in sample one, a mean was calculated across all four raters for each cue so that each target had only one rating per cue. Potential cues were then grouped into valid (detected), valid (undetected), poor, or misleading based on bivariate correlations between the cues and both PLV and actual violent history. Again, valid (detected) cues were correlated with both PLV and violent history, valid (undetected) cues were correlated with only violent history and not PLV, poor cues were not significantly correlated with either PLV or violent history, and misleading cues were correlated with PLV but not actual violent history. It was hypothesized that cues associated with high levels of testosterone (i.e., facial masculinity, heavy brow, apparent physical strength, younger age, and FWHR; e.g., Carré & McCormick, 2008; Salvador et al., 1999; Stillman et al., 2010) and the appearance of mental disorder (e.g., Arseneault et al., 2000; Monahan, 1992; Van Dorn et al., 2011) would be valid cues to violence.

Bivariate correlations were conducted to explore which cues were valid (detected), valid (undetected), poor, and misleading in participants' judgements of violence. Again, cues

indicative of high levels of testosterone, including facial masculinity, heavy brow, apparent physical strength, younger age (e.g., Carré & McCormick, 2008; Salvador et al., 1999; Stillman et al., 2010), and FWHR (e.g., Carré & McCormick, 2008; Carré et al., 2009; Carré et al., 2013; Short et al., 2012; Stillman et al., 2010; Stirrat & Perrett, 2010; Verdonck et al., 1999), were expected to be valid (detected) cues toward violent propensity. Additionally, the appearance of mental disorder was expected to be significantly correlated with both PLV and actual violence (e.g., Arseneault et al., 2000; Monahan, 1992; Van Dorn et al., 2011). Smiling, facial hair, and the emotional expressions of surprise, fear, and sadness (Stillman et al., 2010) were expected to be poor cues, not significantly correlating with either PLV or an actual history of violence. Good grooming, attractiveness, and the facial expressions of disgust, anger, and happiness, were all expected to correlate significantly with PLV but not with an actual history of violence, indicating that they are misleading cues in judgements of violent propensity. Again, ratings of the expression of disgust were left out of the current analyses as they lacked internal consistency. No specific hypotheses were made about the relationships between the exploratory cues, including long hair, facial asymmetry, the absence of glasses, and creepiness with PLV or violent history. These correlations are presented in Table 14 (see Appendix K for correlations among cues, violent history, and participants' PLV ratings by gender).

To reiterate, a family-wise error rate was not used to assess the significance of the results. As research on the relationships between brief exposures to static facial cues and judgements and accuracy of threat detection is a very new area of study, a cutoff value of $p = .05$ was used to determine significance. Consistent with hypotheses and the results from sample one, correlational analyses from sample two identified heaviness of brow (i.e., more violence) and older age (i.e., less violence) as valid (detected) cues to level of violence. Of the exploratory

Table 14

Sample Two: Bivariate Correlations of Facial Cues with PLV and Actual Violent History

Facial Cue	PLV	Violent History
FWHR	.07	-.14
Smile	.17	-.13
Absence of glasses	.15	.22*
Jewelry	-.01	-.14
Facial hair	-.05	-.03
Long hair	.13	.10
Attractiveness	.06	.37***
Creepiness	.24*	-.27*
General facial masculinity	.18	.28*
Heaviness of brow	.36***	.31**
Care in grooming	-.11	.38***
Appearance of physical strength	.31**	.11
Quality of photo	-.31**	-.15
Appearance of physical handicap	-.19	-.25*
Appearance of mental disorder	-.02	-.32**
Facial symmetry	-.19	.31**
Older age	-.40***	-.23*
Appearance of age	-.31**	-.26*
Appearance of arrogance	.25*	.10
Angry expression	.46***	-.12
Fear expression	-.10	.11
Happy expression	-.18	.03
Sad expression	-.01	.15
Surprise expression	-.19	-.05

Note. PLV = Perceived likelihood of violence.

* $p < .05$. ** $p \leq .01$. *** $p \leq .001$.

variables, the appearance of older age was also a valid (detected) cue (i.e., less violence), as it was in sample one. However, the findings from sample one that general facial masculinity and the absence of glasses were correlated with both PLV and history of violence was not replicated in sample two. The appearance of a physical handicap was a valid cue to a history of less violence, but it did not correlate with PLV, making it a valid (undetected) cue to actual violent history. Moreover, creepiness was found to correlate with both PLV and actual history of violence. However, it correlated positively with a PLV ($r = .24, p < .05$) of more violence and

negatively with a history of violence ($r = -.27, p = .01$), suggesting that it served as a misleading cue, leading participants to judge creepy targets as more violent when, in fact, they were less violent.

Consistent with the results from sample one, attractiveness and good grooming were both valid (undetected) cues toward violent propensity, correlating with actual history of violence but not with PLV. Again, this suggests that these cues were valid indicators of a history of violence but were not used by participants in assessments of violent propensity. General facial masculinity was also found to be a valid (undetected) cue to violence, contrary to hypothesis. It was significantly correlated with a history of more violence, as expected, but it was not significantly related to PLV. Moreover, in sample two, this was also found to be true of both facial symmetry and the absence of glasses which were correlated with histories of more violence, but not with participants' ratings of PLV.

Contrary to hypothesis but consistent with the results from sample one, the appearance of a mental disorder was found to be a valid (undetected) cue to a lower propensity for violence. That is, it was not correlated with perceptions of violent propensity but was correlated with a less violent history ($r = -.32, p < .01$). Again, this was also the case for the appearance of a physical handicap.

In terms of poor cues, all of the results from sample one were replicated in sample two. As hypothesized, smiling, facial hair, and the emotional expressions of surprise, fear, and sadness were all poor cues (i.e., no correlations with PLV nor violent history) to violent propensity. Again, the presence of jewelry and the presence of long hair were additional poor cues to violent propensity. Moreover FWHR was a poor cue, contrary to the initial hypothesis that this metric would be a valid indicator of violent propensity but consistent with the findings

from sample one. Again, as in sample one, the expression of happiness was also a poor cue to violence, contrary to the initial prediction that it would be a misleading cue to ratings of less violent propensity.

As hypothesized and consistent with findings from sample one, angry expressions were found to be misleading, relating to the perception of higher propensity for violence. Of the exploratory variables, arrogant appearance was, again, found to be a misleading cue, correlating with perceptions of violence but not with actual violence. Thus, participants seemed to rely on angry and arrogant appearances to erroneously rate targets displaying those cues as having higher inclinations toward violence. Again, photo quality was also found to be a misleading cue. It correlated with level of PLV ($r = -.31, p < .01$) but not with an actual history of violence ($p > .10$), indicating that targets in lower quality photos were rated as having lower propensities for violence than targets in higher quality photos.

The appearance of physical strength, although it was hypothesized to be a valid cue to more violence, was found to be a misleading cue, correlating with higher PLVs but not with actual history of violence. Creepiness was correlated with both PLV and violent history. However, it was technically a misleading cue as it was positively correlated with PLV but negatively correlated with an actual history of violence, suggesting that these targets were rated as more violent even though they were actually less violent. Valid (detected), valid (undetected), poor, and misleading cues from sample two are presented in Table 15.

Of the original 24 facial cues examined, the designations into valid (detected), valid (undetected), poor, and misleading categories were replicated across both samples one and two for 21 cues. The only three which were not replicated were general facial masculinity (valid [detected] in sample one, valid [undetected] in sample two), the absence of glasses (valid

[detected] in sample one, valid [undetected] in sample two), and creepiness (valid [undetected] in sample one, misleading in sample two). A summary of the cue designations (i.e., valid [detected], valid [undetected], poor, and misleading) that were replicated across samples one and two (all participants) is presented in Table 16.

Table 15

Sample Two: Valid, Poor, and Misleading Facial Cues

Valid (Detected)	Valid (Undetected)	Poor	Misleading
Older age *	Absence of glasses	FWHR	Angry expression
Appearance of age*	General facial	Smile	Appearance of
Heaviness of brow	masculinity	Jewelry	physical strength
	Facial symmetry	Facial hair	Appearance of
	Appearance of mental	Long hair	arrogance
	disorder	Happy expression	Quality of photo
	Appearance of	Surprise expression	Creepiness
	physical handicap	Fear expression	
	Care in grooming	Sad expression	
	Attractiveness		

Note. * = cues that were correlated with less propensity for violence. FWHR = facial width-to-height ratio.

Table 16

Valid, Poor, and Misleading Facial Cues Replicated across Samples One and Two

Valid (Detected)	Valid (Undetected)	Poor	Misleading
Older age *	Facial symmetry	FWHR	Angry expression
Appearance of age*	Appearance of mental	Smile	Appearance of
Heaviness of brow	disorder	Jewelry	physical strength
	Appearance of	Facial hair	Appearance of
	physical handicap	Long hair	arrogance
	Care in grooming	Happy expression	Quality of photo
	Attractiveness	Surprise expression	
		Fear expression	
		Sad expression	

Note. * = cues that were correlated with less propensity for violence. FWHR = facial width-to-height ratio.

Valid cues and decision making. As in sample one, a hierarchical binary logistic regression was conducted to examine the contribution of valid (detected) cues to accuracy in distinguishing between violent and less-violent targets. Again, it was expected that valid (detected) cues would contribute to participants' accuracy scores and that poor and misleading cues would not (Stillman et al., 2010). Actual violent history was the dependent variable and PLV was entered as the independent variable in step one of the regression (Wald = 0.33, $p = .57$). Consistent with the results from sample one, PLV was not a significant predictor of actual history of violence (Wald = 0.28, $p = .60$). The cues identified previously as valid for sample two participants (i.e., older age, appearance of older age, and heavy brow) were entered in the next step of this regression. Again, entering these cues did not statistically significantly change the relationship between PLV and violent history (Wald = 0.88, $p = .35$) suggesting that these cues were not used for the assessment of violent propensity. However, as an independent predictor, heaviness of brow was demonstrated to be a statistically significant predictor of violence (Wald = 5.58, $p = .02$).

Three more binary logistic regressions were then conducted to examine whether valid (undetected) cues, poor cues, and misleading cues contributed to the ability to distinguish between violent and less-violent targets. None of these models were found to be good models for the prediction of violence. Moreover, most of the individual cues were not found to be significant predictors of violence, with the exception of creepiness (Wald = 6.12, $p = .01$).

Taken together, the results from sample two are consistent with those from sample one. Participants were not able to distinguish between violent and less-violent targets and, not surprisingly, none of the facial cues contributed to accuracy in such an ability. As in sample one, results from sample two suggest that creepiness, as well as heaviness of brow, may be the only

cues that are predictive of actual violence, but they do not contribute to individuals' abilities to decipher between more and less violent targets.

Psychopathy and accuracy. As described previously, individuals with higher levels of psychopathic personality traits were expected to be more accurate in the detection of violence from facial characteristics than those with lower levels of these traits. This was based on the assumption that psychopathic people are generally more adept in their perceptions of others based on non-verbal cues, as research evidence suggests that these individuals are more accurate in the identification of specific characteristics than those with lower levels of psychopathic traits (e.g., Book et al., 2013; Ritchie, 2016). This specific research question has not yet been examined and, as such, examination of which of the facets of psychopathy, if any, contributed accuracy in these perceptions was exploratory.

First, bivariate correlations were conducted between the facet-level and total scores on measures of psychopathy and participant accuracy scores to determine if level of psychopathic traits was related to accuracy in threat detection (see Table 17). Again, the SRP-4 (Paulhus et al., 2016) and the TriPM (Patrick, 2010) were used as measures of psychopathy for the current study as they are each based on a different conceptualization of psychopathy (e.g., four-facet vs. three-facet, respectively). Results of these correlational analyses demonstrated that, in sample two, only one facet of psychopathy was correlated with accuracy scores – the boldness scale of the TriPM ($r = .15, p < .05$).

All of the facet-level and total scores for both measures of psychopathy were then entered into a multiple regression with accuracy as the dependent variable. According to the F-test, this was not a good predictive model ($R^2 = 0.05, F(9, 123) = 1.80, p = .08$). However, the TriPM boldness ($\beta = -1.412, t(9) = -2.45, p = .01$), meanness ($\beta = 1.62; t(9) = 2.73, p < .01$), disinhibition

Table 17

Sample Two: Bivariate Correlations between Each SRP-4 and TriPM Personality Dimension and Accuracy

Psychopathy Measure	Accuracy
SRP-4	
Interpersonal Manipulation	.05
Callous Affect	.08
Criminal Tendencies	.10
Erratic Lifestyle	.08
SRP-4 Total	.09
TriPM	
Boldness	.15*
Meanness	.04
Disinhibition	.06
TriPM Total	-.02

Note. SRP-4 = Self-Report Psychopathy Scale (Paulhus et al., 2016); TriPM = Triarchic Psychopathy Measure (Patrick, 2010).

* $p < .05$.

($\beta = 1.70$; $t(9) = 2.62$, $p = .01$), and total scores ($\beta = -3.70$; $t(9) = -2.84$, $p < .01$) were all found to be significant predictors in this model. None of the SRP-4 scales were significant predictors in this model (i.e., p values all $> .05$). Considering these findings, it is not surprising that the regression results to determine the predictive ability of solely the SRP-4 facet-level and total scores for accuracy indicated that this regression equation was not significant ($R^2 = 0.01$, $F(5, 149) = 0.36$, $p = .88$) and that none of the individual facet or total scores of the SRP-4 were predictive of accuracy (i.e., all p values $> .05$). However, a multiple regression was also run to determine the roles of only the TriPM facet-level and total scores in the prediction of accuracy. This regression equation was significant ($R^2 = .06$, $F(4, 154) = 2.65$, $p = .04$). Boldness ($\beta = -1.01$; $t(4) = -2.04$, $p = .04$), meanness ($\beta = 1.22$; $t(4) = 2.39$, $p = .02$), disinhibition ($\beta = 1.32$; $t(4) = 2.48$, $p = .01$), and total scores ($\beta = -2.66$; $t(4) = 2.39$, $p = .02$) on the TriPM were all significant predictors of participants' accuracy scores.

General personality and accuracy. To reiterate, no specific a priori hypotheses were made concerning the relationships between the HEXACO personality traits and accuracy in threat detection based on non-verbal facial cues because this has not yet been examined in the literature. To explore these potential relationships, a series of bivariate correlations were first conducted to examine the potential relationships between the HEXACO personality facets and participants' total accuracy (i.e., how often the participants were accurate in violent assessments, ranging from 0-84; see Table 18).

Results from these analyses demonstrated that only the honesty-humility facet of the HEXACO-60 was correlated with participants' accuracy in distinguishing between violent and less-violent offenders ($r = -.18, p = .01$). This finding suggests that participants with lower levels of honesty-humility tended to be more accurate than those with higher levels of this trait. A multiple regression equation was then calculated to explore the potential roles of the HEXACO-60 personality traits in the prediction of accuracy in threat detection based on non-verbal cues. Although this regression equation was insignificant ($R^2 = 0.05, F(6, 180) = 1.47, p = .19$), honesty-humility ($\beta = -0.19; t(6) = -2.26, p = .03$) was found to be a significant predictor of accuracy.

Table 18

Sample Two: Bivariate Correlations between Each HEXACO-60 Personality Dimension and Accuracy

HEXACO-60 Dimension	Accuracy
Honesty-Humility	-.18*
Emotionality	.02
eXtraversion	-.09
Agreeableness	.06
Conscientiousness	-.09
Openness	-.04

* $p < .05$.

Discussion

The purpose of the current investigation was to replicate and extend the findings of Stillman et al. (2010) by contributing novel information to the field of study concerning laypersons' accuracy in person perception based on non-verbal cues. Specifically, one goal of the current study was to determine whether undergraduate students possessed the ability to distinguish between violent and less-violent sex offenders simply by looking at static facial photographs for two seconds. An additional goal was to examine which personality traits, if any, contributed to accuracy in such decision making and, specifically, whether psychopathic personality traits were associated with a greater accuracy in distinguishing between the violent and less-violent offender groups.

Distinguishing Target Violence

The ability of participants to distinguish between violent and less-violent offenders based on two-second exposures to their facial photographs was first examined. Results from the current investigation were inconsistent with Stillman et al.'s (2010) findings in that participants in both sample one and sample two were unable to distinguish between the violent and less-violent groups of offenders. This finding also conflicts with more general findings of surprisingly accurate personality observations based solely on non-verbal cues in the existing literature (e.g., Bar et al., 2006; Carré et al., 2009; Stillman et al., 2010; Willis & Todorov, 2006) and seem to be contradictory to the assumption of a finely-tuned adaptive threat-detection mechanism in humans (e.g., Duntley & Schackelford, 2008).

This result seems to suggest that, in fact, humans are not able to decipher marginal distinctions in threat under such stringent conditions (i.e., minimal exposure to a static image). This may indicate that the ability to perceive threat from non-verbal cues is not a universally

evolved adaptive mechanism that exists among humans. On the other hand, it may be that threat detection is a universal adaptation but that it is not as finely-tuned as hypothesized. As the specific conditions of the current investigation did not involve any dynamic information upon which to base their assumptions, perhaps participants' true threat-detection capabilities were not captured. It may also have been a combination of the information-poor circumstances and the fact that the participants were asked to distinguish only small differences in violence among a sample of known violence (i.e., sex offenders). Perhaps to detect minor distinctions in levels of violent propensity, more dynamic information is required (e.g., video clips; interactions). It is also possible that the utilization of a student sample for ratings of PLV was a hindrance in the current investigation. For example, perhaps there would be a more dramatic, and significant, difference in the perception of violence propensity between these two groups if an offender sample was used, providing a greater range of psychopathy scores among raters. This might better capture an enhanced threat detection ability among individuals with higher levels of psychopathic traits if one exists.

While Stillman et al. (2010) found contradictory results, the sample size of that study ($n = 97$) was small and the effect size of that finding ($d = 0.44$) was small-moderate (Cohen, 1988). The question of power may go some way to help explain the difference in results between Stillman et al.'s (2010) study and the current investigation, in which the finding of no effect was replicated across two large samples.

Gender Differences in Accuracy

Differences in the level of participant accuracy were then examined according to sex. The findings from the current investigation were consistent with previous research and with the assertion that person perception is a universal adaptation in that, across both samples one and

two, males and females were equally accurate in distinguishing between violent and less-violent offenders (e.g., Chan et al., 2010; McArthur & Baron, 1983; Stillman et al., 2010). However, while it was expected that there would be no gender differences in accuracy, it was anticipated that both sexes would be able to distinguish between the violent and less-violent groups, which was not the case. Thus, the finding that males and females were equally accurate seems to suggest, rather, that both genders are equally ineffectual in identifying small differences in levels of violent propensity. Taken together, these results seem to contradict the proposition that an evolutionarily-evolved adaptation exists that is so fine-tuned that humans can distinguish between small differences in violence based merely on very brief exposures to facial cues.

Furthermore, females did not tend to rate the targets as more violent across both groups of offenders, as hypothesized. This finding contrasts those of Stillman et al. (2010) and seems to contradict EMT (i.e., less detrimental mistakes are made in the process of avoiding more detrimental mistakes; Haselton & Buss, 2000), suggesting that females may not be evolutionarily-adapted to be more cautious as a result of greater risks associated with failure to do so. Again, although Stillman et al. (2010) found a contrasting result, their effect size was small-moderate ($d = 0.38$; Cohen, 1988). The results from the current study could suggest the lack of a female-only bias to rate targets as more violent overall. However, it is also possible that such a bias does exist, but only in more dynamic or realistic situations in which females may feel more of an immediate threat to their physical safety. It is well documented empirically that females are more fearful of falling victim to aggression, physical danger, and crime than are males (e.g., Box, Hale, & Andrews, 1988; Campbell, Muncer, & Bibel, 2001; Haghihi & Sorenson, 1996; Harris, 1996; Hill, Howell, & Driver, 1985; Riger, Gordon, & LeBailly, 1982; Vitelli & Endler, 1993). Moreover, females tend to be more risk averse than males (Weber,

Blais, & Betz, 2002) and to perceive a higher level of threat when presented with potentially dangerous situations (e.g., Cheung, Wu, & Tao, 2013; Harris & Miller, 2000). As such, it would be worthwhile for future research to further investigate gender differences in perceived level of threat from more realistic or dynamic situations (e.g., video or audio stimuli) to develop a clearer understanding of this effect.

Valid, Poor, and Misleading Facial Cues

As little evidence exists to clarify the associations between static facial cues and rapid judgements of sex offenders' propensities for violence, one goal of the current study was to investigate such relationships. As such, the facial cues that were utilized, the main hypotheses concerning these cues, and the categorization labels for the current investigation were based on the study conducted by Stillman et al. (2010). Raters who were blind to targets' violence categories first independently rated each of the 84 facial images on a variety of facial cues. Many of these cues have not previously been examined with respect to the specific relationship under investigation in the current study and thus were considered exploratory. The interrater reliability of cue ratings across raters was acceptable for most cues but weaker for several (i.e., creepiness, facial masculinity, facial symmetry, disgust expression, fear expression, sadness expression, and surprise expression). Separate participants then provided ratings for perceived violent propensity (i.e., PLV) for each of the 84 offenders, which were used to examine the associations between PLV and facial cues.

Valid cues. As previously elaborated, many of the facial cues fell into the anticipated categories (i.e., valid [detected], valid [undetected], poor, and misleading). Consistent with previous research, many cues indicative of testosterone, such as general facial masculinity (i.e., more violence), heaviness of brow (i.e., more violence), and older age, as well as appearance of

older age (i.e., less violence), were valid (detected) cues to level of violence for sample one (e.g., Carré & McCormick, 2008; Salvador et al., 1999; Stillman et al., 2010). As previously noted, while the appearance of older age was an exploratory variable in the current investigation, it is not surprising that it was also a valid (detected) cue to less violence, as it was strongly associated with actual age. However, in sample two only three of these facial cues were replicated as valid (detected) cues: older age, the appearance of age, and heaviness of the brow bone. While facial masculinity was also found to be a valid indicator of violent propensity for sample two, it was not related to participants' perceptions of violent propensity. Perhaps this is the case because this cue is more susceptible to alteration than the other three. For instance, it is much easier for a man to shave off his beard, which likely contributes to his overall level of facial masculinity, than it is for him to conceal his age or alter his bone structure. If this is the case, this may suggest that people are more adaptively attuned to more static features that may signal threat than to dynamic features. Thus, perhaps the replicated cues provide more objective measures of an individual's testosterone level.

Oddly, the absence of glasses was a valid (detected) cue to violence for sample one and a valid (undetected) cue for sample two. However, as previously discussed, the relationship between the absence of glasses and a higher likelihood of violence is likely a result of the fact that many more of the less-violent targets were wearing glasses in their mug shots than were the violent targets. Moreover, as the use of this cue by participants in judgements of violence was not replicated in sample two (i.e., this cue was not associated with PLV), it is not likely that the presence or absence of glasses strongly contributes to assessments of an individual's personal propensity toward violence. Again, the inclusion of glasses as a potential facial cue for violent propensity was exploratory as no research exists to date that examines this relationship.

Moreover, most studies examining the influence of facial cues on third party judgements of target characteristics require that targets not have any visible facial adornments, such as glasses (e.g., Freeman, Johnson, Ambady, & Rule, 2010; Kramer, Gottwald, Dixon, & Ward, 2012). However, historically, glasses have been associated with the perception of intelligence (e.g., Saito, 1978; Thornton, 1943, 1944), particularly upon only brief exposure to the targets (Argyle & McHenry, 1971), and lower levels of intelligence quotient (IQ) have been associated with a greater likelihood of violence (e.g., Diamond, Morris, & Barnes, 2012; Rushton & Timpler, 2009). As such, future research in the area of non-verbal cues and threat detection may benefit from a more controlled investigation of the association between the presence of glasses and third parties' perceptions of violence likelihood.

Another unexpected finding was that both attractiveness and good grooming, which were anticipated to be cues that would lead participants to erroneously rate targets as less violent, were actually valid (undetected) cues toward violence. This finding suggests that these cues were not actually used by participants in forming their perceptions about violence but that they were, in fact, related to an actual history of violence. Although inconsistent with previous research indicating that individuals' judgements are influenced by attractiveness (e.g., Capetillo et al., 2008; Castellow et al., 1990; Curno, 2015; Esses & Webster, 1988; Langlois et al., 2000) and suggesting a link between good grooming and perceptions of less violence (Stillman et al., 2010), this finding was replicated across both samples one and two. It is possible that this unanticipated finding relates to the associations between attractiveness, testosterone, and violence. Specifically, targets' levels of grooming were strongly associated with their levels of attractiveness, which have previously been associated with testosterone levels (e.g., Roney et al., 2006). As illustrated in more detail above, higher levels of testosterone have been associated

with more violence and aggression (e.g., Salvador et al., 1999; Carré & McCormick, 2008). Thus, while participants did not rely on either of these cues to inform their perceptions of violence, they may actually serve as valid indicators of a greater inclination toward violence. This may be the case particularly for attractiveness, as care in grooming may fluctuate from day to day. One potential reason that participants did not appear to use these cues may relate to the fact that all of the targets used in the current investigation were sex offenders. Future research should examine whether this finding replicates in a study using other types of violent individuals as targets (i.e., not only sex offenders).

Facial symmetry is another, more static, cue that has been previously linked to attractiveness (e.g., Bashour, 2006; Borelli & Berneburg, 2010; Grammer & Thornhill, 1994; Penton-Voak, Jones, Little, Baker, Tiddeman, Burt, & Perrett, 2001; Perrett, Burt, Penton-Voak, Lee, Rowland, & Edwards, 1999). Thus, given the associations highlighted above, it is not surprising that, across both samples one and two, this exploratory variable was also found to be a valid cue for an inclination toward violence. However, facial symmetry was not used by participants in their ratings of violent propensity in either sample. Scheib, Gangestad, and Thornhill (1999) found similar results. Specifically, they found that there was an association between participants' ratings of target attractiveness and targets' facial symmetry but that participants were not actually detecting symmetry, since this relationship remained even after symmetry cues were removed (i.e., only one half of the face was displayed). Moreover, their participants were unable to accurately rate the symmetry of target faces. These results suggest that other cues were being relied upon by participants in their evaluations of attractiveness (Scheib et al., 1999), indicating that a third variable may be responsible for the association between attractiveness and symmetry. For example, in terms of the current investigation, perhaps

level of testosterone is related to both attractiveness and facial symmetry and it is by this relationship that these variables are associated with violence. The nature of the associations that exist among attractiveness, testosterone, facial symmetry, and violence remain somewhat unclear as the data are correlational. Thus, there may be additional, unmeasured variables associated with facial symmetry that explain the association between facial symmetry and target violence.

Consistent with expectations, the appearance of a mental disorder was associated with targets' actual history of violence. However, contrary to expectations and to the existing evidence concerning the association between mental disorders and violence (e.g., Arseneault et al., 2000; Monahan, 1992; Van Dorn et al., 2011), the appearance of a mental disorder was related to a target history of less violence. Moreover, counter to the initial hypothesis based on the literature concerning the public perception of mental disorder and violence (e.g., Field Institute, 1984; Monahan, 1992), this cue went undetected by participants. This was the case across both samples in the current investigation. Interestingly, this was also the case for the appearance of a physical handicap, which was included in the current investigation as an exploratory variable.

Most of the information that influences public perceptions of mental disorder comes from the media (Economou, Stefanis, & Papadimitriou, 2009) and negative attitudes about mental illness prevail (Rose, Thornicroft, Pinfold, & Kassam, 2007; Wahl, 2003). However, some research suggests that educational programs can be successful in temporarily reducing negative attitudes and eradicating negative stereotypes about mental illness (Corrigan et al., 2002; Penn, Kommana, Mansfield, & Link, 1999; Sakellari, Leino-Kilpi, & Kalokerinou-Anagnostopoulou, 2011). Perhaps the fact that participants did not seem to rely on the appearance of mental disorder relates to an increased societal focus on abolishing stigma surrounding mental disorder.

That is, it may be that in the setting of an institution for higher education (i.e., a university), students have had more exposure to educational platforms concerning mental disorders, and even physical disorders, leading to fewer negative attitudes and decreased stigmatization. This could potentially explain why participants did not seem to rely on such cues in making their decisions about violence.

Poor cues. Poor cues, which were not related to either targets' actual violent history or participants' judgements of violence, were the same across both samples. Consistent with hypotheses and previous research (i.e., Stillman et al., 2010), smiling, facial hair, and the emotional expressions of surprise, fear, and sadness all fell into this category. Moreover, among the exploratory variables included in the current study, jewelry and long hair were also found to be poor cues to violent propensity. However, contrary to prediction, FWHR and the emotional expression of happiness were also found to be poor cues. As stated previously, FWHR is a measure of masculinity that has not only been associated with aggression but has also been found to play a role in individuals' assessments of aggression in target individuals (e.g., Carré & McCormick, 2008; Carré et al., 2009; Lefevre & Lewis, 2014; Short et al., 2012; Stirrat & Perrett, 2010). As such, it was hypothesized that a larger FWHR would provide a valid cue to a higher level of actual violent history. Unfortunately, the photographs used for the current investigation were not standardized and, as such, did not lend themselves well to the accurate measurement of target FWHR. A difference as simple as how the photograph is taken (e.g., Třebický, Fialová, Kleisner, & Havlíček, 2016) or how the target is positioned (e.g., Hehman, Leitner, & Gaertner, 2013) can affect the measurement of this metric and, as such, the validity of the measure used in the current studies is questionable. Based on the results of Stillman et al. (2010), it was also expected that the emotional expression of happiness would be a misleading

cue, leading participants to rate targets as less violent. However, the expression of happiness was also considered to be a poor cue in terms of violent proclivity. This finding was replicated across both samples, suggesting that expressions of happiness may not be as misleading in judgements of violence as anticipated.

Misleading cues. The same cues that were found to be misleading in sample one were also found to be misleading in sample two. This lends credence to the assumption that these cues truly serve to erroneously influence individuals' judgements of others' levels of violence. Moreover, the existence of misleading cues in the current investigation is consistent with DDT which, again, suggests that individuals often make decisions about the characteristics of other people based on lasting first impressions formed using cues which can often be misleading (Porter & ten Brinke, 2009). As anticipated, and consistent with the finding of Stillman et al. (2010), expressions of anger were related to participants' judgements of more violence, even when the targets were less-violent. Of the exploratory variables in the current investigation, photo quality, the appearance of arrogance, and creepiness were all found to be misleading cues. It is interesting and perhaps not surprising to note that angry and arrogant expressions are both cues that may lead individuals to arbitrarily assume that another individual is more violent than they actually are. The fact that expressions of arrogance may inform inaccurate judgements of more violence is a novel finding and, as such, further replication is needed beyond the two samples in the current investigation to conclude that it truly influences judgements of violence.

The finding that photo quality influenced participants' perceptions of violence is another important finding from the current investigation. Specifically, offenders with lower quality mugshots were rated as having a lower perceived propensity for violence than those with higher quality mugshots. This finding further demonstrates that extraneous aspects of target

photographs can have a great impact on a study's findings (e.g., Třebický et al., 2016) and thus highlights the importance of uniform, standardized photographs for studies investigating the use of cues in person perception. This may also be important to note for use in applied settings. For example, this highlights the importance of ensuring standardized, high-quality photographs for use in photo lineups. Future research may wish to examine this relationship in more detail to determine whether photo quality influences decision making in legal contexts such as this.

Interestingly, in the current study, creepier targets tended to be less-violent. That is, less-creepy individuals tended to have a more violent history and, as such, this was a valid cue. In sample one, this cue went undetected by participants. However, in sample two, creepier targets were often rated as more violent. In this sense, this suggested that creepiness is actually a misleading cue, as individuals perceived less violent individuals as more violent. It is possible that the inconsistency between the two samples relates to the fact that consistency of the ratings of creepiness among the four raters fell below an acceptable level. However, this discrepancy was marginal and likely would not have had a great influence on the quality of the rating. As creepiness was included in the current investigation as an exploratory variable, replication in future studies of first impressions of violence based on facial cues is necessary in order to further clarify this potential association.

Based on the fact that testosterone is associated with aggression and violence (e.g., Carré & McCormick, 2008; Salvador et al., 1999), it was originally hypothesized that a number of testosterone-related cues (i.e., heaviness of brow, facial masculinity, younger age, appearance of physical strength) would be valid indicators to participants' perceptions of targets' level of violence. However, contrary to this prediction, physical strength was found to be a misleading cue for both samples, associated with higher perceptions of target violence but not with targets'

actual histories of violence. This finding is also inconsistent with those of Stillman et al. (2010), who found that physical strength was, in fact, a valid cue to target violence. Stillman et al. (2010) did, however, make note of the fact that the appearance of physical strength was more strongly correlated with PLV than with actual violent history, suggesting that participants may have overused this cue in their sample. It is possible that the use of larger samples in the current investigation allowed for a more precise evaluation of the use of this cue in violence detection.

Valid Cues and Decision Making

It was hypothesized that, consistent with Stillman et al. (2010), valid cues identified in the previous analyses would contribute to participants' accuracy in distinguishing between violent and less-violent targets. However, contrary to expectation, none of the valid (detected) cues contributed to participant accuracy in either sample one or sample two. That is, these cues did not significantly improve participants' accuracy in decision-making about target violence. This result is not surprising considering the finding that participants in both samples were not able to distinguish between violent and less-violent targets. However, the results do suggest that a few of these cues may be predictors of actual violence, without contributing to an ability to actually decipher between more and less violent individuals, such as heaviness of brow and creepiness. In fact, the finding that creepiness was a significant independent predictor of actual violent history was replicated across both samples. It is unclear why individuals who appear more objectively creepy may actually be less violent individuals. To examine this more thoroughly, one might begin to parse out what aspects of an individual's appearance lead that person to be rated as more or less creepy. This may be an interesting avenue for future investigation. Future research may wish to examine what factors contribute to the 'creepiness' of

a person, as it is unclear what exactly it is about a person that makes them appear creepy and how that might relate to their level of violence.

Psychopathy and Accuracy

The current study was the first to explore the relationship between psychopathic personality traits and accuracy in first impression judgements of threat, based solely on static facial features. As it has previously been found that individuals with higher levels of psychopathic traits are more successful in detecting vulnerability from others' non-verbal cues (e.g., Book et al., 2013; Ritchie, 2016), it was hypothesized that these individuals may also exhibit increased accuracy in the detection of threat. Such a result may suggest that psychopathic individuals are generally more adept in perceiving the characteristics of others based on non-verbal cues than individuals with lower levels of psychopathic traits. The specific relationships between each of the facets of the SRP and the TriPM and accuracy in the detection of violent propensity has not yet been explored and, as such, no specific hypotheses were made about these relationships.

Upon examination of the predictive values of psychopathic traits on accuracy in violence detection in sample one, the callous affect (CA) facet of the SRP-4 was the only facet, on either the SRP or the TriPM, which had some predictive value. This was only the case when all psychopathic traits and total scores from both the three-factor TriPM and four-factor SRP were examined simultaneously. When just the facet and total scores of the SRP were examined, callous affect was not predictive of accuracy. This may suggest that, although all together these traits do not contribute to an ability to detect an actual history of violence, callous affect might provide some advantage in this ability. To determine whether this is true, future research should further explore this relationship to determine whether these findings are replicated.

However, the lack of a predictive relationship could have been a result of the fact that the self-report measures were not displayed in randomized order in sample one. Fatigue may have influenced the results in this sample by contributing to increased attrition and/or missing responses pooled mostly on the last survey(s). Moreover, there was only one attention check (i.e., quality control) question within these questionnaires in sample one, which may not have been sufficient to ensure accurate responding by participants throughout the surveys. Thus, more confidence may be placed in the results of sample two, for which more highly controlled methods were implemented (i.e., more quality control questions and randomized questionnaire order).

When all psychopathic traits and total scores from both measures were examined for predictive ability, this combination was, again, not predictive of accuracy. However, all factors of the three-factor TriPM model – boldness, meanness, and disinhibition, as well as the total score for the TriPM – were independently predictive of accuracy. Specifically, higher levels of boldness and of total psychopathy on the TriPM were predictive of greater accuracy and lower levels of meanness and disinhibition were predictive of greater accuracy. Moreover, when the predictive ability of this combination of psychopathic traits was examined without the traits of the SRP, it was found that it predicted participants' accuracy scores. That is, individuals with higher levels of boldness and total psychopathy on the TriPM and individuals with lower levels of meanness and disinhibition on the TriPM were more accurate in deciphering between more and less violent targets. It is possible that including more quality control questions and randomizing the order of the personality questionnaires for sample two allowed for more accurate measures of psychopathy and personality (i.e., controlling for fatigue effects and attrition) resulting in a change of results from those of sample one. It would seem that, despite

the debate in the literature concerning the factor structure of the TriPM (e.g., Crego & Widiger, 2014; Lilienfeld et al., 2012; Lilienfeld et al., 2016; Lynam & Miller, 2012; Miller & Lynam, 2012; Neumann et al., 2013; Patrick et al., 2013), there is something about this combination of personality traits that played a role in the detection of violence in this sample. These results suggest that this factor structure may better capture the personality traits that contribute to psychopathic individuals' increased accuracy in person perception, at the very least when it comes to the detection of violence. Specifically, boldness and total psychopathy as measured by the TriPM may capture personality traits that contribute to accuracy in the detection of violence propensity in others. Additionally, the results suggest that individuals who are less mean and those who are less disinhibited may also be better able to make accurate decisions about threat. However, replication of these findings in future research, and across various types of samples, is needed before it can be firmly concluded that these traits enhance accuracy in threat detection.

General Personality and Accuracy

In sample one, the findings from analyses examining the potential relationships between the HEXACO personality traits and accuracy in the detection of actual violence were similar to those for psychopathic personality traits and such accuracy. Specifically, increased accuracy was not related to differing levels on any of the HEXACO dimensions. Moreover, not surprisingly, examination of the predictive value of these personality dimensions revealed that none were predictive of accuracy in the detection of violence. These results suggest that no specific general personality traits contribute to individuals' ability to accurately distinguish between subtle variations in levels of violence.

However, in study two the honesty-humility dimension of the HEXACO was found to be related to accuracy, such that individuals with lower levels of honesty-humility were more

accurate in their perceptions of violence. Furthermore, although the HEXACO dimensions taken together were not found to be predictive of accuracy in perceptions of violence, the honesty-humility dimension was. This suggests that individuals who scored lower on this dimension were more successful in distinguishing between small differences in levels of violence than those with higher honesty-humility. This finding is very interesting as lower levels of honesty-humility have been found to be associated with higher levels of psychopathy (e.g., Lee & Ashton, 2005; Lee et al., 2013; Ritchie, 2016). This finding, taken together with the findings that higher boldness and total psychopathy on the TriPM were significant predictors of greater accuracy in the current sample, suggest that there may be some merit to the hypothesis that individuals with higher levels of psychopathic traits tend to be more accurate in threat detection based solely on brief glimpses at static, non-verbal cues.

Implications

The results of the current study do not seem to support the evolutionary perspective of a threat detection adaptation in humans (Duntley & Schackelford, 2008). Neither the results of sample one nor sample two replicated the existing research findings that human beings, regardless of gender, are able to detect threat based on very minimal information (e.g., Carré et al., 2009; Stillman et al., 2010). However, as previously stated, this may only be the case in situations in which (a) dynamic information is not available, (b) the difference in level of violence is subtle, or (c) both dynamic information is not available and the difference in level of violence is subtle. As previously noted, the current investigation was a very stringent test of this ability and, as such, it may be premature to wholly discount this perspective.

The results of the current investigation expand upon the existing body of literature that examines human beings' abilities to automatically detect characteristics of others based on brief

exposure to static, non-verbal cues. They provide insight into the ability of humans to immediately perceive non-verbal cues and to use that information to make decisions about threat posed by others. The majority of this research thus far has centred its focus on what *target* features contribute to participants' target evaluations. This is the case for Stillman et al.'s (2010) study examining the ability of laypeople to detect violence from faces. The current study however, also focused on the personality characteristics of the *raters* that may have been associated with a greater accuracy in assessing others' propensity for violence from facial characteristics. Moreover, the results from the current investigation contribute important information to the limited existing literature regarding which non-verbal cues are attended to in the perception of violence in others

In terms of personality characteristics of the raters, the findings from the current study may begin to answer the question of whether psychopathic individuals are more accurate in person perception in general, or whether this ability is specific to the detection of vulnerability in others. The findings above provide new evidence that individuals with higher levels of certain psychopathic traits may also be better able to detect threat based on non-verbal cues. In particular, higher levels of boldness as measured by the TriPM, may tap into the specific characteristics of psychopathy that enhance an ability to quickly detect threat, as higher levels of TriPM boldness and total psychopathy were predictive of increased accuracy. This may be suggestive of a generally heightened perception of non-verbal cues among psychopathic individuals, allowing these people to be more accurate in the perception of a wider range of characteristics in others. Alternatively, this heightened ability may not extend beyond the detection of vulnerability and threat from non-verbal cues.

The relationship between psychopathy and accuracy in violence detection contributes novel information to the research domain focused on examining evolutionary adaptations for the avoidance of danger. The current results suggest that the ability to detect threat in others is not a universal evolutionary adaptation that exists in all humans. However, they do provide some evidence to suggest that there are certain personality characteristics that contribute to accuracy in threat detection. This may provide support for the notion that certain personality characteristics in humans developed as evolutionary adaptations to enhance survival. For example, it has been proposed that psychopathy is an evolutionary adaptation that evolved to allow these individuals to be more successful in identifying targets for exploitation (e.g., Glenn et al., 2011; Krupp et al., 2012). If the results from the current investigation are replicated in various populations, this may provide support for the supposition that psychopathy is also an adaptation for threat avoidance. In particular, high boldness may have provided some adaptive advantage to these individuals throughout evolutionary time. Alternatively, as previously suggested, it is possible that part of this ability to more successfully detect vulnerable others may be the ability to also detect who is not vulnerable or, more specifically, who is actually a threat to their own safety. If this is the case, this has the potential to contribute new and valuable information to research on the victim selection process of psychopathic individuals.

In terms of cues to violence, the replication of many cue categorizations (i.e., as valid [detected], valid [undetected], poor, and misleading) across both samples in the current investigation, particularly when consistent with Stillman et al.'s (2010) designations, provides strong evidence for the roles of these cues in humans' decision making about threat. Specific knowledge about which of these cues are valid indicators to violence and which are not is critical for a thorough understanding of decision-making errors about threat. Awareness and cognizance

of the cues that lead individuals to falsely assume people are not dangerous (e.g., appearance of physical strength, photo quality), for instance, may be useful in a number of contexts, such as in legal decision-making situations. However, as this is a relatively new area of investigation, the results from the current investigation must be replicated across samples from different populations before any specific real-world conclusions can be confidently made.

Limitations

Much like any empirical investigation, the current study was subject to a number of limitations. First and foremost were those associated with the study administration to sample one. First, due to human error, only one quality control question (i.e., attention filter) was used throughout the administration of all three self-report personality questionnaires. This may not have provided sufficient control to ensure that all participants included in analyses were paying adequate attention to the task at hand to provide accurate responses. Moreover, in this sample, the questionnaire order was not randomized. This may have presented an issue as the survey was 45-minutes long and was administered online. The lack of randomization could have contributed to potential order effects or fatigue effects in sample one. However, these administration concerns were rectified for sample two, in which one quality control question was randomly placed within each questionnaire and the questionnaires were administered in randomized order to control for order effects and or/fatigue effects. The difference in results concerning personality between sample one and sample two may indicate that these factors are a concern and that the results of sample one should be interpreted with caution.

Other measurement limitations apply to both samples. For instance, with self-report data, there is always the concern that participants are susceptible to a social-desirability bias. This may be of particular concern for participants' responses on measures of psychopathic traits, as

psychopathy is a sensitive construct. Self-reported responses become increasingly more susceptible to socially-desirable responding with increased sensitivity of the construct (Tourangeau, Rips, & Rasinski, 2000). However, evidence suggests that self-report psychopathy measures are valid and that data collected thereby are not significantly affected by response distortion (Watts, Lilienfeld, Edens, Douglas, Skeem, & LoPilato, 2016). Research also suggests that, in comparison with interviews, using a self-report methodology tends to decrease participants' levels of inaccurate responding (Fowler, 1995; Tourangeau & Smith, 1996, 1998; Tourangeau & Yan, 2007) which, theoretically, is a result of the privacy associated with such reports (Dalal & Hakel, 2016). Moreover, the self-report measures for the current study were administered on an encrypted online platform, which ensured the anonymity of participants' responses. Participants were made aware of the fact that all responses would be collected through a server that uses multiple layers of security to protect the privacy of data and that their responses would be coded in such a way that they could not possibly be identified. However, online administration presents some limitations of its own in terms of control. For instance, it is difficult to know whether participants paid adequate attention throughout the administration of the study (e.g., whether they looked at the photos for the full two seconds).

The fact that five items on the HEXACO-60 were unintentionally missing from the online survey might be another factor that presents a limitation for the current study. However, each missing item was from a separate dimension. Thus, this problem was easily rectified by statistical means. A notable statistical limitation of the current study does exist, however, with respect to the cue ratings used for analyses. Several of the cues ratings that were included in the analyses did not meet an acceptable inter-rater consistency level. Acknowledging this, however, the majority of these cues did not fall dramatically below the reported acceptable value and those

which did were removed from analyses. Nonetheless, results concerning the use of these cues should be interpreted with caution. Moreover, one limitation of any correlational study is the lack of experimental manipulation. Without this, conclusions about cause and effect cannot be made. This is another reason to be cautious about the interpretations that can be drawn from the cue designations in the current investigation.

Another limitation of the current study was the use of non-standardized photos. As discussed previously, the photos used were the same mug shots used by Stillman et al. (2010) that were extracted from a public online sex offender registry. As such, it was not possible to control for differences in photo characteristics (e.g., zoom, focus, pixellation, proportions, etc.). As such, these photos did not lend themselves well to the measurement of participants' FWHR. This may explain the lack of an association between this variable and either PLV by participants or actual violent histories of the targets. It is also unknown whether this lack of control over the uniformity of the photos influenced any other photo ratings. Another potential measurement concern may relate to the use of public records of the targets' offences as the measure of actual violent history. This may present a problem because, while this provides a true measure of real-life violence rather than an experimentally-construed measure of violence, it does not account for potential acts of violence committed for which the target was not caught and arrested.

The nature of the current investigation produces some limitations to the external validity of the study. For instance, the study design itself made use of a manufactured environment in which people only had two seconds to look at a still picture of a face, with no dynamic information, and make a rapid judgement about violence likelihood. Thus, the results of the current investigation may not be completely reflective of real-world situations in which detection of threat is needed (e.g., dark alley, legal decision making). For instance, perhaps individuals

would be more successful in deciphering between violent and non-violent individuals in situations that provide more dynamic information to the observer. However, the purpose of the current study was to examine this ability in situations lacking dynamic information and, specifically, whether this ability is fine-tuned enough to make these decisions based only on first impressions of faces, with no extraneous information.

A further limitation to the external validity of the current investigation is the use of undergraduate students as the sample. As it was found that the factors of the TriPM contributed to participants' accuracy in sample two, it may be possible that, if we had more highly scoring 'psychopaths' in our sample, a significant difference in PLV between the violent and less-violent groups would have been found. Thus, these findings may not generalize to an observer sample of, for example, offenders – a population with higher rates of psychopathy than the general population – when presented with the same task. Moreover, university students compose a unique population of individuals with distinct demographic characteristics (e.g., higher IQ, middle-class, etc.) and the individuals who willingly volunteer to participate in research studies such as this one are inherently different from those who do not. Both of these factors may limit the ability to generalize the current results to the general population or to other populations (e.g., offenders). Specifically, it is possible that a more even distribution of HEXACO personality characteristics exists in the general population and that the current sample did not capture a sufficient range to detect differences in accuracy across these dimensions. In addition to these limitations, the majority of participants in each sample was female. This may have limited the capacity of the current study to capture true gender differences.

Future Directions

In attempting to replicate the findings from the current investigation (i.e., with respect to cue designations and particularly the associations between honesty-humility and TriPM scores and accuracy), future research may seek to account for the limitations described above. For example, the collection and use of standardized photographs of offenders as targets, rather than using mug shots, may allow for more accurate measurement of certain facial features. Further, the collection of violent history by self-report from the target offenders, in addition to using their public record information, may allow for more comprehensive measurement of their actual violent history.

Future researchers may also wish to expand on the current findings in a number of ways. For instance, the inclusion of a strictly non-violent group of target individuals may provide unique information about the potential existence of a threat detection adaptation. That is, with less subtle differences between the levels of violence of the different target groups, perhaps participants would be able to successfully distinguish between them. Such a finding may indicate that a threat detection adaptation in humans does exist, but that it may not be as finely-tuned as anticipated. A further extension along this line of investigation may be to conduct a very similar study, but use target photographs of individuals who are not sex offenders. This would allow for an examination of any potential changes in perceptions of targets between a known sex offender sample and a sample of 'regular people'.

Another way in which future research may wish to expand upon the findings from the current investigation is to examine the effect of more dynamic cues in this context. Examination of whether laypeople are able to distinguish between sex offenders' levels of violent history based on very brief glimpses of dynamic pieces of information, such as brief video clips, may

reveal that this information is necessary to detect such subtle differences. Researchers may also examine what effect this has on the relationships between participants' personality characteristics, including psychopathic, and their accuracy. Moreover, collection of reports from participants about what cues they used in their decision making would provide stronger evidence for cue designations and could provide additional information about any unmeasured cues that are often used in these decisions.

Another interesting avenue for expansion on this line of investigation would be to use an incarcerated sample as participants. As there are higher levels of psychopathy in prison populations (i.e., 15-25%; Hemphill, Hare, & Wong, 1998; Salekin, Rogers, & Sewell, 1996) than in the general population (i.e., < 1%; Coid, Yang, Ullrich, Roberts, & Hare, 2009; Hare, 1998; Neumann & Hare, 2008; Ullrich, Farrington, & Coid, 2008) a stronger examination of the link between psychopathic traits and success in the identification of violent individuals based on minimal information may be facilitated. If the link between psychopathic personality characteristics and accuracy found in sample two is replicated – specifically, the predictive relationships between higher TriPM boldness and total psychopathy and accuracy – perhaps a sample of offenders would be better able to distinguish the very subtle differences in level of violence between the violent and less-violent targets based only on brief exposure to their faces.

Finally, it would be prudent for future research to conduct further investigation into what an enhanced threat detection ability in psychopathic individuals would truly mean. For instance, would this indicate a generally enhanced ability in psychopathic individuals to accurately perceive the characteristics of other people? Conversely, does this ability only exist with respect to the detection of vulnerability and threat? If the latter is the case, further examination may be warranted in order to decipher whether or not these enhanced abilities are part of the same

mechanism. Results suggesting that they are may provide important information about the victim selection process of psychopathic individuals.

Conclusion

Overall, this study did not provide evidence for the existence of a universal threat detection adaptation in all humans. However, the roles of a number of facial cues in decision making about threat were replicated across both samples. Some of the cue designations found by Stillman et al. (2010) were replicated in the current investigation and some were not. As such, further replication using more consistent ratings of facial cues is required to adequately understand their roles in decision making about threat. Moreover, although there was inconsistency between the results of samples one and two with respect to personality differences in accuracy, the more controlled methodology implemented in sample two produced results suggestive of an enhanced violence-detection ability in individuals with higher levels of certain psychopathic traits. Specifically, the current investigation produced the first piece of evidence to suggest that the boldness factor and total psychopathy scores of the TriPM model of psychopathy, as well as low scores on the honesty-humility dimension of the HEXACO, may provide individuals with an enhanced ability to detect actual violence in others. While it is essential that these findings be replicated across other samples, the results presented above do indicate that the relationship between psychopathic personality traits and accuracy in threat detection warrants further investigation.

References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality & Social Psychology*, 55(3), 387-395.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York, NY: Holt.
- Altman, D. G. (1991). Mathematics for kappa. In D. Altman (Ed.). *Practical statistics for medical research* (pp. 406-407). London, UK: Chapman & Hall.
- Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 201-271). San Diego, CA: Academic Press.
- Ambady, N., Hallahan, M., & Conner, B. (1999). Accuracy of judgements of sexual orientation from thin slices of behavior. *Journal of Personality and Social Psychology*, 77(3), 538-547.
- Ambady, N., Krabbenhoft, M. A., & Hogan, D. (2006). The 30-sec sale: Using thin-slice judgments to evaluate sales effectiveness. *Journal of Consumer Psychology*, 16(1), 4-13. doi:10.1207/s15327663jcp1601_2
- Ambady, N., & Rosenthal, R. (1992). Thin slices of behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, 2, 256-274.
- Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, 64(3), 431-441. doi:10.1037/0022-3514.64.3.431
- Archer, J. (2006). Testosterone and human aggression: An evaluation of the challenge hypothesis. *Neuroscience and Biobehavioral Reviews*, 30, 319-345.

doi:10.1016/j.neubiorev.2004.12.007.

Argyle, M., & McHenry, R. (1971). Do spectacles really affect judgments of intelligence?

British Journal of Social and Clinical Psychology, 10, 27–29.

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, 11(2), 150-166.

Ashton, M. C., & Lee, K. (2008). The prediction of honesty-humility-related criteria by the HEXACO and Five-Factor Models of personality. *Journal of Research in Personality*, 42, 1216-1228. doi: 10.1016/j.jp. 2008.03.006

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, 91, 340-345.

doi:10.1080/00223890902935878.

Ashton, M. C., Lee, K., & de Vries, R. E. (2014). The HEXACO Honesty-Humility, Agreeableness, and Emotionality factors: A review of research and theory. *Personality and Social Psychology Review*, 18, 139-152.

Ashton, M. C., Lee, K., & Goldberg, L. R. (2004a). A hierarchical analysis of 1,710 English personality-descriptive adjectives. *Journal of Personality and Social Psychology*, 87(5), 707-721. doi: 10.1037/0022-3514.87.5.707

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., De Vries, R. E., Di Blas, L., ... De Raad, B. (2004b). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, 86, 356-366.

Ashton, M. C., Lee, K., & Son, C. (2000). Honesty as the sixth factor of personality: Correlations

- with Machiavellianism, primary psychopathy, and social adroitness. *Europe Journal of Personality*, 14, 359-369.
- Arseneault, L., Moffitt, T. E., Caspi, A., Taylor, P. J., & Silva, P. A. (2000). Mental disorders and violence in a total birth cohort. *Archives of General Psychiatry*, 57, 979-986.
- Babiak, P., Neumann, C. S., & Hare, R. D. (2010). Corporate psychopathy: Talking the walk. *Behavioral Sciences & the Law*, 28, 174-193. doi: 10.1002/bsl.925
- Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, 6(2), 269-278.
doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/1528-3542.6.2.269
- Bashour, M. (2006). History and current concepts in the analysis of facial attractiveness. *Plastic Reconstructive Surgery*, 118(3), 741-756.
- Black, P. J., Woodworth, M., & Porter, S. (2014). The Big Bad Wolf? The relation between the Dark Triad and the interpersonal assessment of vulnerability. *Personality and Individual Differences*, 67, 52-56. doi:10.1016/j.paid.2013.10.026
- Blair, R. J. R. (2010). Psychopathy, frustration, and reactive aggression: The role of ventromedial prefrontal cortex. *British Journal of Psychology*, 101(3), 383-399.
- Bohl, V., & Bos, W. V. (2012). Toward an integrative account of social cognition: Marrying theory of mind and interactionism to study the interplay of Type 1 and Type 2 processes. *Frontiers in Human Neuroscience*, 6(274), 1-15. doi:10.3389/fnhum.2012.00274
- Book, A. S., Costello, K., & Camilleri, J. A. (2013). Psychopathy and victim selection: The use of gait as a cue to vulnerability. *Journal of Interpersonal Violence*, 23, 2386-2383.
doi:10.1177/088620512475315
- Book, A. S., Quinsey, V. L., & Langford, D. (2007). Psychopathy and the perception of affect and vulnerability. *Criminal Justice and Behavior*, 34, 531-544.

doi:10.1177/0093854806293554

- Borelli, C., & Berneburg, M. (2010). "Beauty lies in the eye of the beholder"? Aspects of beauty and attractiveness. *Journal of the German Society of Dermatology*, 8(5), 326-330.
- Box, S., Hale, C., & Andrews, G. (1988). Explaining fear of crime. *British Journal of Criminology*, 28, 340-356.
- Camilleri, J. A., Kuhlmeier, V. A., & Chu, J. Y. Y. (2010). Remembering helpers and hinderers depends on behavioural intentions of the agent and psychopathic characteristics of the observer. *Evolutionary Psychology*, 8(2), 303-316.
- Campbell, A., Muncer, S., & Bibel, D. (2001). Women and crime: An evolutionary approach. *Aggression and Violent Behavior*, 6, 481-497.
- Campbell, A., Muncer, S., & Bibel, D. (2001). Women and crime: An evolutionary approach. *Aggression and Violent Behavior*, 6, 481-497.
- Capetillo, J. (2008). *Gender differences in attribution and person perception* (Unpublished doctoral dissertation) Vanderbilt University, Nashville, TN.
- Carlston, D. E., & Skowronski, J. J. (2005). Linking versus thinking: Evidence for the different associative and attributional bases of spontaneous trait transference and spontaneous trait inference. *Journal of Personality and Social Psychology*, 89(6), 884-898.
- Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B: Biological Sciences*, 275(1651), 2651-2656.
- doi:10.1098/rspb.2008.0873
- Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial Structure Is a Reliable Cue of Aggressive Behavior. *Psychological Science*, 20(10), 1194-1198. doi:10.1111/j.1467-

9280.2009.02423.x

- Carré, J. M., Morrissey, M. D., Mondloch, C. J., & McCormick, C. M. (2010). Estimating aggression from emotionally neutral faces: Which cues are diagnostic? *Perception*, 39(3), 356-377. Doi: 10.1068/p6543
- Carré, J. M., Murphy, K. R., & Hariri, A. R. (2011). What lies beneath the face of aggression? *Social Cognitive and Affective Neuroscience*, 8(2), 224-229. doi:10.1093/scan/nsr096
- Castellow, W. A., Wuensch, K. L., & Moore, C. H. (1990). Effects of physical attractiveness of the plaintiff and defendant in sexual harassment judgments. *Journal Social Behavior and Personality*, 5, 547-562.
- Chan, M., Rogers, K. H., Parisotto, K. L., & Biesanz, J. C. (2011). Forming first impressions: The role of gender and normative accuracy in personality perception. *Journal of Research in Personality*, 45(1), 117-120. doi:10.1016/j.jrp.2010.11.001
- Cheung, H. Y., Wu, J., & Tao, J. (2013). Risk perception and risk-taking attitude: A comparison between Hong Kong and Mainland Chinese undergraduate students. *The Asia-Pacific Education Researcher*, 22(4), 497-506. doi:10.1007/s40299-012-0048-7
- Cleckley, H. (1941). *The mask of sanity: An attempt to clarify some issues about the so-called psychopathic personality*. St. Louis, MO: Mosby.
- Clifford, S., & Jerit, J. (2014). Is there a cost to convenience? An experimental comparison of data quality in laboratory and online studies. *Journal of Experimental Political Science*, 1, 120-131. doi:10.1017/xps.2014.5
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Coid, J., Yang, M., Ullrich, S., Roberts, A., & Hare, R. D. (2009). Prevalence and correlates of

- psychopathic traits in the household population of Great Britain. *International Journal of Law and Psychiatry*, 32, 65-73. doi: 10.1016/j.ijlp.2009.01.002
- Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment*, 13(2), 171-188.
doi:<http://dx.doi.org.proxy.library.carleton.ca/10.1037/1040-3590.13.2.171>
- Cooke, D. J., Michie, C., Hart, S. D., & Clark, D. (2005). Searching for the pan-cultural core of psychopathic personality disorder. *Personality and Individual Differences*, 39, 283-295. doi: 10.1016/j.paid.2005.01.004
- Corr, P. J. (2010). The psychoticism–psychopathy continuum: A neuropsychological model of core deficits. *Personality and Individual Differences*, 48(6), 695–703.
- Corrigan, P. W., Rowan, D., Green, A., Lundin, R., River, P., Uphoff-Wasowski, K., . . . Kubiak, M. A. (2002). Challenging two mental illness stigmas: Personal responsibility and dangerousness. *Schizophrenia Bulletin*, 28(2), 293-309.
doi:10.1093/oxfordjournals.schbul.a006939
- Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.
- Crego, C., & Widiger, T. A. (2014). Psychopathy, DSM–5, and a caution. *Personality Disorders: Theory, Research, and Treatment*, 5, 335–347. <http://dx.doi.org/10.1037/per0000078>
- Curno, J. (2015). *Risky business: The role of psychopathy and violence risk assessment in forensic decision-making* (Unpublished master's thesis). Carleton University, Ottawa, ON.

- Dalal, D. K., & Hakel, M. D. (2016). Experimental comparisons of methods for reducing deliberate distortions to self-report measures of sensitive constructs. *Organizational Research Methods*, 19(3), 475-505. doi:10.1177/1094428116639131
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York, NY: Putnam.
- Degroot, T., & Gooty, J. (2009). Can nonverbal cues be used to make meaningful personality attributions in employment interviews? *Journal of Business Psychology*, 24, 179-192.
- Depaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, 111(2), 203-243. doi:10.1037/0033-2909.111.2.203
- de Waal, F. (2007). *Chimpanzee politics: Power and sex among apes*. Baltimore, MD: Johns Hopkins University Press.
- Diamond, B., Morris, R. G., & Barnes, J. (2012). Individual and group IQ predict inmate violence. *Intelligence*, 40(2), 115-122. doi:10.1016/j.intell.2012.01.010
- Duntley, J. D., & Shackelford, T. K. (2008). *Evolutionary forensic psychology: Darwinian foundations of crime and law* (pp. 201-229). New York, NY: Oxford University Press.
- Economou, M., Stefanis, N. C., & Papadimitriou, G. N. (2009). Schizophrenia and stigma: Old problems, new challenges. In S. Kasper & G. N. Papadimitriou (Eds.), *Schizophrenia: Biopsychosocial approaches and current challenges* (pp. 299-309). London: Informa Healthcare.
- Edens, J. F., Marcus, D. K., Lilienfeld, S. O., & Poythress, N. G. (2005). Psychopathic, not psychopath: Taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology*, 115(1), 131-144. doi: 10.1037/0021-843X.115.1.131
- Esses, V. M., & Webster, C. D. (1988). Physical attractiveness, dangerousness, and the Canadian

- criminal code. *Journal of Applied Social Psychology*, 18, 1017–1031.
- Faul, F., Erdfelder, E., Lang, A-G., & Buckner, A. (2009). G*power (Version 3.1) [Computer Software]. Germany: University of Dusseldorf.
- Field, A. P. (2013). *Discovering statistics using IBM SPSS* (4th ed.). London: SAGE.
- Field Institute. (1984). *In pursuit of wellness: A survey of California adults*. Sacramento, CA: California Department of Mental Health.
- Fowler, F. J. (1995). *Improving survey questions: Design and evaluation*. Thousand Oaks, CA: SAGE Publications.
- Fowler, K.A., Lilienfeld, S. O., & Patrick, C.J. (2009). Detecting psychopathy from thin slices of behavior. *Psychological Assessment*, 21(1), 68-78.
- Frank, R. H. (1988). *Passions within reason: The strategic role of emotions*. New York, NY: Norton.
- Freeman, J. B., Johnson, K. L., Ambady, N., & Rule, N. O. (2010). Sexual orientation perception involves gendered facial cues. *Personality and Social Psychology Bulletin*, 36(10), 1318-1331. doi:10.1177/0146167210378755
- Frith, C. D., & Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, 60(3), 503-510. doi:10.1016/j.neuron.2008.10.032
- Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin*, 101(1), 75-90.
- Furley, P., & Schweitzer, G. (2016). In a flash: Thin slice judgment accuracy of leading and trailing in sports. *Journal of Nonverbal Behavior*, 40(2), 83-100. doi:10.1007/s10919-015-0225-5
- Glenn, A. L., Kurzban, R., & Raine, A. (2011). Evolutionary theory and psychopathy.

- Aggression and Violent Behavior*, 16(5), 371-380. doi:10.1016/j.avb.2011.03.009
- Goffman, E. (1979). *Gender advertisements*. New York, NY: Harper & Row.
- Gillen, C. T., Bergström, H., & Forth, A. E. (2016). Individual differences and rating errors in first impressions of psychopathy. *Evolutionary Psychology*, 14, 1-7.
- Goodwin, B. E. (2015). *A nomological network of psychopathy: Elaborating on the construct validity of the triarchic model* (Unpublished master's thesis), University of Alabama, Tuscaloosa, AL.
- Grammer, K., & Thornhill, R. (1994). Human (homo sapiens) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology*, 108(3), 233-242.
doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0735-7036.108.3.233
- Grayson, B., & Stein, M. I. (1981). Attracting assault: Victims' nonverbal cues. *Journal of Communications*, 31, 68-75. doi:10.1111/j.1460-2466.1981.tb01206.x
- Green, D. M., & Swets, J. A. (1966). *Signal detection and psychophysics*. New York, NY: Wiley.
- Haghihi, B., & Sorenson, J. (1996). America's fear of crime. In T. J. Flanagan & D. R. Longmire (Eds.), *Americans view crime and justice: A national public opinion survey* (pp. 16-30). Thousand Oaks: Sage Publications.
- Hall, J. (1984). *Nonverbal sex differences*. Baltimore: Johns Hopkins University Press.
- Hare, R. D. (1991). *Manual for the revised psychopathy checklist* (1st ed.). Toronto, ON: Multi-Health Systems.
- Hare, R. D. (1998). The Hare PCL-R: Some issues concerning its use and misuse. *Legal and Criminological Psychology*, 3, 99-119.

- Hare, R. D. (2001). Psychopaths and their nature: Some implications for understanding human predatory violence. In A. Raine & J. Sanmartin (Eds.), *Violence and psychopathy* (pp. 5-34). New York, NY: Kluwer Academic/Plenum.
- Hare, R. D. (2003). *Manual for the Hare Psychopathy Checklist-Revised* (2nd ed.). Toronto, ON: Multi-Health Systems.
- Harpending, H. C., & Sobus, J. (1987). Sociopathy as an adaptation. *Ethology and Sociobiology*, 8(1), 63–72.
- Harris, M. B. (1996). Aggression, gender, and ethnicity. *Aggression and Violent Behavior*, 1(2), 123–146.
- Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: Facial structure predicts unethical behaviour. *Proceedings of the Royal Society B*, 279(1728), 571–576.
- Haselton, M. G., Buss, D. M., & DeKay, W. T. (1998, July). *A theory of errors in cross-sex mindreading*. Paper presented at the Human Behavior and Evolution Society Meeting, Davis, CA.
- Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, 78(1), 81–91.
- Helman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology*, 49(4), 747-754.
doi:10.1016/j.jesp.2013.02.015
- Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology*, 3, 139-170.
- Hill, G. D., Howell, F. M., & Driver, E. T. (1985). Gender, fear, and protective handgun ownership. *Criminology*, 23, 541–552.

- Jones, E. E. (1990). *Interpersonal perception*. New York, NY: Freeman.
- Jones, D. N., & Paulhus, D. L. (2011). The role of impulsivity in the dark triad of personality. *Personality and Individual Differences*, 51(5), 679-682.
doi:10.1016/j.paid.2011.04.011
- Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin*, 102(3), 390-402.
- Kosson, D. S., Neumann, C. S., Forth, A. E., Salekin, R. T., Hare, R. D., Krischer, M. K., & Sevecke, K. (2013). Factor structure of the Hare Psychopathy Checklist: Youth Version (PCL:YV) in adolescent females. *Psychological Assessment*, 25, 71-83. doi: 10.1037/a0028986
- Kohn, A. R., & Fiedler, R. E. (1961). Age and sex differences in the perceptions of persons. *Sociometry*, 24(2), 157-164.
- Koppensteiner, M., Stephan, P., & Jäschke, J. P. M. (2016). Moving speeches: Dominance, trustworthiness and competence in body motion. *Personality and Individual Differences*, 94, 101-106.
- Kramer, R. S., Gottwald, V. M., Dixon, T. A., & Ward, R. (2012). Different cues of personality and health from the face and gait of women. *Evolutionary Psychology*, 10(2), 271-295.
doi:10.1177/147470491201000208
- Krupp, D. B., Sewall, L. A., Lalumière, M. L., Sheriff, C., & Harris, G. T. (2012). Nepotistic patterns of violent psychopathy: Evidence for adaptation? *Frontiers in Psychology*, 3.
doi:10.3389/fpsyg.2012.00305
- Lambert, N. M, Mulder, S., & Fincham, F. (2014). Thin slices of infidelity: Determining whether observers can pick out cheaters from a video clip interaction and what tips them off.

- Personal Relationships*, 21, 612-619.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analysis and theoretical review. *Psychological Bulletin*, 126, 390-423.
- Lee, K. & Ashton, M. C. (2004). Psychometric properties of the HEXACO Personality Inventory. *Multivariate Behavioural Research*, 39, 329-358. doi: 10.1207/s15327906mbr3902_8.
- Lee, K., & Ashton, M. C. (2005). Psychopathy, Machiavellianism, and narcissism in the five-factor model and the HEXACO model of personality structure. *Personality and Individual Differences*, 38, 1571-1582. doi:10.1016/j.paid.2004.09.016
- Lee, K., & Ashton, M. C. (2008). The HEXACO Personality Factors in indigenous personality lexicons of English and 11 other languages. *Journal of Personality*, 76(5), 1001-1054.
- Lee, K., Ashton, M. C., Wiltshire, J., Bourdage, J. S., Visser, B. A., & Galucci, A. (2013). Sex, power, and money: Prediction from the Dark Triad and honesty-humility. *European Journal of Personality*, 27, 169-184. doi:10.1002/per.1860
- Leedom, L. J., & Almas, L. H. (2012). Is psychopathy a disorder or an adaptation? *Frontiers in Psychology*, 3. doi:10.3389/fpsyg.2012.00549
- Lefevre, C. E., & Lewis, G. J. (2014). Perceiving aggression from facial structure: Further evidence for a positive association with facial width-to-height ratio and masculinity, but not for moderation by self-reported dominance. *European Journal of Personality*, 28(6), 530-537. doi:10.1002/per.1942
- Leistico, A. R., Salekin, R., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and Human*

- Behaviour*, 32, 28-45. doi: 10.1007/s10979-007-9096-6.
- Lilienfeld, S. O., Patrick, C. J., Benning, S. D., Berg, J., Sellbom, M., & Edens, J. F. (2012). The role of fearless dominance in psychopathy: Confusions, controversies, and clarifications. *Personality Disorders: Theory, Research, and Treatment*, 3, 327–340.
<http://dx.doi.org/10.1037/a0026987>
- Lilienfeld, S. O., Smith, S. F., Sauvigné, K. C., Patrick, C. J., Drislane, L. E., Latzman, R. D., & Krueger, R. F. (2016). Is boldness relevant to psychopathic personality? Meta-analytic relations with non-psychopathy checklist-based measures of psychopathy. *Psychological Assessment*, 28(10), 1172-1185.
[doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/pas0000244](http://dx.doi.org.proxy.library.carleton.ca/10.1037/pas0000244)
- Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198-1202.
[doi:10.2307/2290157](https://doi.org/10.2307/2290157)
- Lynam, D. R., & Miller, J. D. (2012). Fearless dominance and psychopathy: A response to Lilienfeld et al. *Personality Disorders: Theory, Research, and Treatment*, 3, 341–353.
<http://dx.doi.org/10.1037/a0028296>
- Maner, J. K., Kenrick, D. T., Neuberg, S. L., Becker, D. V., Robertson, T., Hofer, B., ..., Schaller, M. (2005). Functional projection: How fundamental social motives can bias interpersonal perception. *Journal of Personality and Social Psychology*, 88, 63-78.
- Martelli, M., Majib, J. M., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, 5, 58-70.
- McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, 90(3), 215-238. doi:10.1037/0033-295x.90.3.215

- McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the Five-Factor Model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), 81-90.
- Mealey, L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. *Behavioral and Brain Sciences*, 18, 523-599. doi:10.1017/S0140525X 00039595
- Miller, J. D., & Lynam, D. R. (2012). An examination of the Psychopathic Personality Inventory's nomological network: A meta-analytic review. *Personality Disorders: Theory, Research, and Practice*, 3, 305–326. <http://dx.doi.org/10.1037/a0024567>
- Miller, A. K., Rufino, K. A., Boccaccini, M.T., Jackson, R. L., & Murrie, D. C. (2011). On individual differences in person perception: Raters' personality traits relate to their psychopathy checklist–revised scoring tendencies. *Assessment*, 18(2), 253-260.
- Monahan, J. (1992). Mental disorder and violent behavior: Perceptions and evidence. *American Psychologist*, 47(4), 511-521.
- Murzynski, J., & Degelman, D. (1996). Body language of women and judgments of vulnerability to sexual assault. *Journal of Applied Social Psychology*, 26(18), 1617-1626. doi: 10.1111/j.1559-1816.1996.tb00088.x
- Neumann, C. S., & Hare, R. D. (2008). Psychotic traits in a large community sample: Links to violence, alcohol use, and intelligence. *Journal of Consulting and Clinical Psychology*, 76, 893-899. doi:10.1037/0022-006X.76.5.893
- Neumann, C. S., Kosson, D. S., & Salekin, R. T. (2007). Exploratory and confirmatory factor analysis of the psychopathy construct: Methodological and conceptual issues. In C. S. Neumann, D. S. Kosson, & R. T. Salekin (Eds.), *The psychopath: Theory, research, and practice* (pp. 79-104). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

- Neumann, C. S., Uzieblo, K., Crombez, G., & Hare, R. D. (2013). Understanding the Psychopathic Personality Inventory (PPI) in terms of the unidimensionality, orthogonality, and construct validity of PPI-I and -II. *Personality Disorders: Theory, Research, and Treatment*, 4, 77–79. <http://dx.doi.org/10.1037/a0027196>
- Newman, J. P., Curtin, J. J., Bertsch, J. D., & Baskin-Sommers, A. R. (2010). Attention moderates the fearlessness of psychopathic offenders. *Biological Psychiatry*, 67(1), 66–70.
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32), 11087-11092. doi:10.1073/pnas.0805664105
- Patrick, C. J. (2010). *Operationalizing the Triarchic conceptualization of psychopathy: Preliminary description of brief scales for assessment of boldness, meanness, and disinhibition*. Unpublished manuscript, Florida State University, Tallahassee, FL.
- Patrick, C. J., Venables, N. C., & Drislane, L. E. (2013). The role of fearless dominance in differentiating psychopathy from antisocial personality disorder: Comment on Marcus, Fulton, and Edens. *Personality Disorders: Theory, Research, and Treatment*, 4, 80 – 82. <http://dx.doi.org/10.1037/a0027173>
- Paulhus, D. L., Neumann, C. S., & Hare, R. D. (2016). *Self-Report Psychopathy Scale: Technical manual*. Toronto, ON: Multi-Health Systems.
- Penn, D. L., Kommana, S., Mansfield, M., & Link, B. G. (1999). Dispelling the stigma of schizophrenia II: The impact of information on dangerousness. *Schizophrenia Bulletin*, 25, 437-446.
- Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior*, 25(4), 229-241.

doi:10.1016/j.evolhumbehav.2004.04.003

Penton-Voak, I. S., Jones, B. C., Little, A. C., Baker, S., Tiddeman, B., Burt, D. M., & Perrett, D.

I. (2001). Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. *Proceedings of the Royal Society B: Biological Sciences*, 268(1476), 1617–1623. <http://doi.org/10.1098/rspb.2001.1703>

Perrett, D. I., Burt, D., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999).

Symmetry and human facial attractiveness. *Evolution and Human Behavior*, 20(5), 295–307. doi:10.1016/s1090-5138(99)00014-8

Porter, S., & ten Brinke, L. (2009). Dangerous decisions: A theoretical framework for

understanding how judges assess credibility in the courtroom. *Legal and Criminological Psychology*, 14, 119–134.

Pretsch, J., Flunger, B., Heckmann, N., & Schmitt, M. (2013). Done in 60 s? Inferring teachers'

subjective well-being from thin slices of nonverbal behavior. *Social Psychology of Education*, 16(3), 421–434. doi:10.1007/s11218-013-9223-9

Richards, L., Rollerson, B., & Phillips, J. (1991). Perceptions of submissiveness: Implications for

victimization. *The Journal of Psychology*, 125(4), 407–411. doi: 10.1080/00223980.1991.10543302

Riger, S., Gordon, M. T., & LeBailly, R. K. (1982). Coping with urban crime: Women's use of

precautionary behaviors. *American Journal of Community Psychology*, 10, 369–386.

Ritchie, M. B. (2014). *Perceptions of victim vulnerability: The role of personality and*

psychopathy (Unpublished undergraduate thesis). Carleton University, Ottawa, ON.

Ritchie, M. B. (2016). *Telltale gait: The role of nonverbal cues in perceptions of personality and*

vulnerability (Unpublished master's thesis). Carleton University, Ottawa, ON.

- Roney, J. R., Hanson, K. N., Durante, K. M., & Maestripieri, D. (2006). Reading men's faces: Women's mate attractiveness judgments track men's testosterone and interest in infants. *Proceedings of the Royal Society*, 273, 2169-2175. doi:10.1098/rspb.2006.3569
- Rose, D., Thornicroft, G., Pinfold, V., & Kassam, A. (2007). 250 labels used to stigmatize people with mental illness. *BMC Health Services Research*, 7, 97.
- Rubin, D.B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: John Wiley & Sons, Inc.
- Rule, N. O., Adams, R. B., Ambady, N., & Freeman, J. B. (2012). Perceptions of dominance following glimpses of faces and bodies. *Perception*, 41, 687-706.
- Rushton, J. P., & Templer, D. I. (2009). National differences in intelligence, crime, income, and skin color. *Intelligence*, 37(4), 341-346. doi:10.1016/j.intell.2009.04.003
- Saito, K. (1978). An experimental study of personality judgment: Effect of wearing glasses. *Japanese Journal of Experimental Social Psychology*, 17, 121-127.
- Sakellari, E., Leino-Kilpi, H., & Kalokerinou-Anagnostopoulou, A. (2011). Educational interventions in secondary education aiming to affect pupils' attitudes toward mental illness: A review of the literature. *Journal of Psychiatric and Mental Health Nursing*, 18, 166-176.
- Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the Psychopathy Checklist and the Psychopathy Checklist-Revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice*, 3(3), 203-215. doi:10.1111/j.1468-2850.1996.tb00071.x
- Salvador, A., Suay, F., Martinez-Sanchis, S., Simon, V. M., & Brain, P. F. (1999). Correlating testosterone and fighting in male participants in judo contests. *Physiology and Behavior*,

- 68(1-2), 205-209. doi:10.1016/s0031-9384(99)00168-7
- Schaller, M. (2008). Evolutionary bases of first impressions. In N. Ambady & J. J. Skowronski (Eds.), *First impressions* (pp.1-15). New York, NY: Guilford Press.
- Schmid Mast, M., & Hall, J. A. (2004). Who is the boss and who is not? Accuracy of judging status. *Journal of Nonverbal Behavior*, 29(3), 145-165.
- Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society B: Biological Sciences*, 266(1431), 1913-1917. doi:10.1098/rspb.1999.0866
- Sell, A., Cosmides, L., Tooby, J., Sznycer, D., Rueden, C. V., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society B: Biological Sciences*, 276(1656), 575-584. doi:10.1098/rspb.2008.1177
- Short, L. A., Mondloch, C. J., McCormick, C. M., Carré, J. M., Ma, R., Fu, G., & Lee, K. (2012). Detection of propensity for aggression based on facial structure irrespective of face race. *Evolution and Human Behavior*, 33(2), 121-129. doi: 10.1016/j.evolhumbehav.2011.07.002
- Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, 9(4), 1-11.
- Stillman, T. F., Maner, J. K., & Baumeister, R. F. (2010). A thin slice of violence: Distinguishing violent from nonviolent sex offenders at a glance. *Evolution and Human Behavior*, 31(4), 298-303. doi:10.1016/j.evolhumbehav.2009.12.001
- Stirrat, M., & Perrett, D. (2010). Valid facial cues to cooperation and trust: Male facial width

- and trustworthiness. *Psychological Science*, 21(3), 349-354. doi: 10.1177/0956797610362647
- Swerdlloff, R. S., & Wang, C. (1993). Androgen deficiency and aging. *The Western Journal of Medicine*, 159, 579-585.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55. Doi: 10.5116/ijme.4dfb.8dfd
- Thornton, G. R. (1943). The effect upon judgments of personality of varying a single factor in a photograph. *Journal of Social Psychology*, 28, 203-207.
- Thornton, G. R. (1944). The effect of wearing glasses upon judgments of personality traits of persons seen briefly. *Journal of Applied Social Psychology*, 28, 203-207.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. New York, NY: Cambridge University Press.
- Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, 60, 275-304.
- Tourangeau, R., & Smith, T. W. (1998). Collecting sensitive information with different modes of data collection. In M. P. Couper, R. P. Baker, J. Bethlehem, C. Z. F. Clark, J. Martin, W. L. Nicholls, II, & J. M. O'Reilly (Eds.), *Computer assisted survey information collection* (pp. 431-454). New York, NY: John Wiley & Sons.
- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133, 859-883.

- Třebický, V., Fialová, J., Kleisner, K., & Havlíček, J. (2016). Focal length affects depicted shape and perception of facial images. *PLoS ONE*, *11*(2), 1-14.
<http://doi.org.proxy.library.carleton.ca/10.1371/journal.pone.0149313>
- Ullrich, S., Farrington, D. P., & Coid, J. W. (2008). Psychopathic personality traits and life success. *Personality and Individual Differences*, *44*, 1162-1171. doi:
10.1016/j.paid.2007.11.008
- Valentova, J., Rieger, G., Havlicek, J., Linsenmeier, J., & Bailey, J. (2011). Judgements of sexual orientation and masculinity-femininity based on thin slices of behavior: A cross-cultural comparison. *Archives of Sexual Behavior*, *40*(6), 1145-1152.
- Van Dorn, R., Volavka, J., & Johnson, N. (2012). Mental disorder and violence: is there a relationship beyond substance use? *Social Psychiatry*, *47*(3), 487-503.
- Verdonck, A., Gaethofs, M., Carels, C., & de Zegher, F. (1999). Effect of low-dose testosterone treatment on craniofacial growth in boys with delayed puberty. *European Journal of Orthodontics*, *2*, 137-143, DOI: 0.1093/ejo/21.2.137.
- Vitelli, R., & Endler, N. S. (1993). Psychological determinants of fear of crime: A comparison of general and situational prediction models. *Personality and Individual Differences*, *14*, 77-85.
- Wahl, O. (2003). Depictions of mental illnesses in children's media. *Journal of Mental Health*, *12*, 249-258.
- Watts, A. L., Lilienfeld, S. O., Edens, J. F., Douglas, K. S., Skeem, J. L., Verschuere, B., & LoPilato, A. C. (2016). Does response distortion statistically affect the relations between self-report psychopathy measures and external criteria? *Psychological Assessment*, *28*(3), 294-306. doi:<http://dx.doi.org.proxy.library.carleton.ca/10.1037/pas0000168>

- Weber, E.U., Blais, A., & Betz, N.E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*, 15, 263–290.
- Wheeler, S., Book, A., & Costello, K. (2009). Psychopathic traits and perceptions of victim vulnerability. *Criminal Justice and Behavior*, 36(6), 635-648. doi: 10.1177/0093854809333958
- Willis, J., & Todorov, A. (2006). First Impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17(7), 592-598. doi:10.1111/j.1467-9280.2006.01750.x
- Yuan, Y. C. (2010). Multiple imputation for missing data: Concepts and new development (Version 9.0). *SAS Institute Incorporated*, 49, 1-11.

Appendix A**SONA Recruitment Bulletin**

NAME OF RESEARCHERS: Mari Shanahan Somerville

FACULTY SUPERVISOR: Adelle Forth

RESEARCHER CONTACT INFORMATION:

EMAIL: MariShanahanSomervil@cmail.carleton.ca

FACULTY ADVISOR: adelle.forth@carleton.ca; 613-520-2600 ext. 1267

TITLE OF STUDY: Spot the Violent Sex Offender: Identifying Violent and Less-violent Sex Offenders at a Glance

BRIEF DESCRIPTION: In this study, we are examining accuracy in the identification of violence from just a short glance at a facial photograph. You will be asked to rate a series of photographs based on your perception of their inclination toward violence. In addition, you will be asked to fill out a number of questionnaires designed to assess your personality characteristics (e.g., psychopathic traits, dominance, extraversion). This study will take approximately 45 minutes to complete and you will receive 0.50% credit for your participation. This study has been approved by the Carleton University Research Ethics Board – B.

IS THIS A WEB BASED STUDY? Yes.

IS THIS A TWO-PART STUDY? No.

LENGTH OF STUDY: 45 minutes.

SELECTION CRITERIA: None.

ETHICS CLEARANCE NUMBER (REB #): 105574

Appendix B

Sample Sex Offender Photographs

Nonviolent Sex Offenders

Violent Sex Offenders

Appendix C**Rating Scales of Non-Verbal Cues**

Rater ID.

What is your age?

What is your racial/ethnic background?

- Aboriginal Canadian/Native Canadian/First Nations
- Asian
- Black/African-Canadian
- East Indian
- Hispanic/Latino
- Middle Eastern
- White/Caucasian
- Other: _____

Is the person in this photograph smiling?

- ☐ Yes
☐ No

Is the person in this photograph wearing glasses?

- ☐ Yes
☐ No

Is the person in this photograph wearing visible jewellery?

- ☐ Yes
☐ No

Does the person in this photograph have a visible tattoo?

- ☐ Yes
☐ No

Does the person in this photograph have facial hair?

- ☐ Yes
☐ No

Does the person in this photograph have long hair (i.e., below the ears)?

- ☐ Yes
☐ No

What is the background colour in this photograph?

- ☐ Blue
☐ Gray

Rate the photographs on the degree to which the following apply. Please refer back to the photograph when answering each question.

How attractive is the person in this photograph?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
 (9)

How creepy-looking is the person in this photograph?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
 (9)

In general, how masculine does this person's face appear?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
 (9)

How heavy is the brow (bone structure) of the person in this photograph?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
 (9)

Does it appear that the person in this photograph takes care in their personal grooming?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
 (9)

Does the person in this photograph exhibit physical strength?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
 (9)

How clear is the quality of the photograph itself?

Not at all *Very*

(1) (2) (3) (4) (5) (6) (7) (8) (9)

Does the person in this photograph appear to be handicapped?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

Does the person in this photograph appear to have a mental disorder?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

How symmetrical is the face of the person in this photograph?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

How old does the person in this photograph appear to be?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

How arrogant does the person in this photograph appear?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

Rate the degree to which the person in the photograph is displaying the following facial expressions. Please refer back to the photograph when answering each question.

Is the person in this photograph displaying an expression indicative of anger?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

Is the person in this photograph displaying an expression indicative of disgust?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

Is the person in this photograph displaying an expression indicative of fear?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

Is the person in this photograph displaying an expression indicative of happiness?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

Is the person in this photograph displaying an expression indicative of sadness?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

Is the person in this photograph displaying an expression indicative of surprise?

Not at all (1) (2) (3) (4) (5) (6) (7) (8) *Very*
(9)

Appendix D**Demographic Questionnaire**

Note: You may skip any questions you do not wish to answer.

1. Age: _____
2. Are you...
 - a. Male
 - b. Female
 - c. Other: _____
3. What is your racial/ethnic background?
 - a. Aboriginal Canadian/Native Canadian/First Nations
 - b. Asian
 - c. Black/African-Canadian
 - d. East Indian
 - e. Hispanic/Latino
 - f. Middle Eastern
 - g. White/Caucasian
 - h. Other: _____

Appendix E**Attention Filters****Sample One Attention Filter**

1. Select 2 when answering this question.

1 2 3 4 5 6

Sample Two Attention Filters

1. Select 2 disagree when answering this question.

1 2 3 4 5

2. Select 5 agree strongly when answering this question.

1 2 3 4 5

3. If you are a student at Carleton University, select 3 somewhat false.

1 2 3 4

Appendix F**HEXACO-60**

On the following pages you will find a series of statements about you. Please read each statement and decide how much you agree or disagree with that statement. Then write your response in the space next to the statement using the following scale:

- 5 = Strongly Agree
4 = Agree
3 = Neutral (neither agree nor disagree)
2 = Disagree
1 = Strongly Disagree

Please answer every statement, even if you are not completely sure of your response.

- 1 I would be quite bored by a visit to an art gallery.
- 2 I plan ahead and organize things, to avoid scrambling at the last minute.
- 3 I rarely hold a grudge, even against people who have badly wronged me.
- 4 I feel reasonably satisfied with myself overall.
- 5 I would feel afraid if I had to travel in bad weather conditions.
- 6 I wouldn't use flattery to get a raise or promotion at work, even if I thought it would succeed.
- 7 I'm interested in learning about the history and politics of other countries.
- 8 I often push myself very hard when trying to achieve a goal.
- 9 People sometimes tell me that I am too critical of others.
- 10 I rarely express my opinions in group meetings.
- 11 I sometimes can't help worrying about little things.
- 12 If I knew that I could never get caught, I would be willing to steal a million dollars.
- 13 I would enjoy creating a work of art, such as a novel, a song, or a painting.
- 14 When working on something, I don't pay much attention to small details.
- 15 People sometimes tell me that I'm too stubborn.
- 16 I prefer jobs that involve active social interaction to those that involve working alone.
- 17 When I suffer from a painful experience, I need someone to make me feel comfortable.
- 18 Having a lot of money is not especially important to me.
- 19 I think that paying attention to radical ideas is a waste of time.
- 20 I make decisions based on the feeling of the moment rather than on careful thought.
- 21 People think of me as someone who has a quick temper.

- 22 _____ On most days, I feel cheerful and optimistic.
- 23 _____ I feel like crying when I see other people crying.
- 24 _____ I think that I am entitled to more respect than the average person is.
- 25 _____ If I had the opportunity, I would like to attend a classical music concert.
- 26 _____ When working, I sometimes have difficulties due to being disorganized.
- 27 _____ My attitude toward people who have treated me badly is “forgive and forget”.
- 28 _____ I feel that I am an unpopular person.
- 29 _____ When it comes to physical danger, I am very fearful.
- 30 _____ If I want something from someone, I will laugh at that person's worst jokes.
- 31 _____ I’ve never really enjoyed looking through an encyclopedia.
- 32 _____ I do only the minimum amount of work needed to get by.
- 33 _____ I tend to be lenient in judging other people.
- 34 _____ In social situations, I’m usually the one who makes the first move.
- 35 _____ I worry a lot less than most people do.
- 36 _____ I would never accept a bribe, even if it were very large.
- 37 _____ People have often told me that I have a good imagination.
- 38 _____ I always try to be accurate in my work, even at the expense of time.
- 39 _____ I am usually quite flexible in my opinions when people disagree with me.
- 40 _____ The first thing that I always do in a new place is to make friends.
- 41 _____ I can handle difficult situations without needing emotional support from anyone else.
- 42 _____ I would get a lot of pleasure from owning expensive luxury goods.
- 43 _____ I like people who have unconventional views.
- 44 _____ I make a lot of mistakes because I don’t think before I act.
- 45 _____ Most people tend to get angry more quickly than I do.
- 46 _____ Most people are more upbeat and dynamic than I generally am.
- 47 _____ I feel strong emotions when someone close to me is going away for a long time.
- 48 _____ I want people to know that I am an important person of high status.
- 49 _____ I don’t think of myself as the artistic or creative type.
- 50 _____ People often call me a perfectionist.
- 51 _____ Even when people make a lot of mistakes, I rarely say anything negative.
- 52 _____ I sometimes feel that I am a worthless person.

- 53 _____ Even in an emergency I wouldn't feel like panicking.
- 54 _____ I wouldn't pretend to like someone just to get that person to do favors for me.
- 55 _____ I find it boring to discuss philosophy.
- 56 _____ I prefer to do whatever comes to mind, rather than stick to a plan.
- 57 _____ When people tell me that I'm wrong, my first reaction is to argue with them.
- 58 _____ When I'm in a group of people, I'm often the one who speaks on behalf of the group.
- 59 _____ I remain unemotional even in situations where most people get very sentimental.
- 60 _____ I'd be tempted to use counterfeit money, if I were sure I could get away with it.

Appendix G

Brief TriPM

Directions: This questionnaire contains statements that different people might use to describe themselves.

Each statement is followed by four choices:

1 = True 2 = somewhat true 3 = somewhat false 4 = False

For each statement, fill in the bubble for the choice that describes you best. There are no right or wrong answers; just choose the answer that best describes you.

Like this: ☒ Not like this: ☒ ☒ ☒

Remember: Fill only one bubble per item. If you make a mistake cross out the incorrect answer with an X and fill in the correct option. Answer all of the items. Please work rapidly and do not spend too much time on any one statement.

1. I'm optimistic more often than not.	1	2	3	4
2. How other people feel is important to me.	1	2	3	4
3. I often act on immediate needs.	1	2	3	4
4. I have no strong desire to parachute out of an airplane.	1	2	3	4
5. I've often missed things I promised to attend.	1	2	3	4
6. I would enjoy being in a high-speed chase.	1	2	3	4
7. I am well-equipped to deal with stress.	1	2	3	4
8. I don't mind if someone I dislike gets hurt.	1	2	3	4
9. My impulsive decisions have caused problems with loved ones.	1	2	3	4
10. I get scared easily.	1	2	3	4
11. I sympathize with others' problems.	1	2	3	4
12. I have missed work without bothering to call in.	1	2	3	4
13. I'm a born leader.	1	2	3	4

14.	I enjoy a good physical fight.	1	2	3	4
15.	I jump into things without thinking.	1	2	3	4
16.	I have a hard time making things turn out the way I want.	1	2	3	4
17.	I return insults.	1	2	3	4
18.	I've gotten in trouble because I missed too much school.	1	2	3	4
19.	I have a knack for influencing people.	1	2	3	4
20.	It doesn't bother me to see someone else in pain.	1	2	3	4
21.	I have good control over myself.	1	2	3	4
22.	I function well in new situations, even when unprepared.	1	2	3	4
23.	I enjoy pushing people around sometimes.	1	2	3	4
24.	I have taken money from someone's purse or wallet without asking.	1	2	3	4
25.	I don't think of myself as talented.	1	2	3	4
26.	I taunt people just to stir things up.	1	2	3	4
27.	People often abuse my trust.	1	2	3	4
28.	I'm afraid of far fewer things than most people.	1	2	3	4
29.	I don't see any point in worrying if what I do hurts someone else.	1	2	3	4
30.	I keep appointments I make.	1	2	3	4
31.	I often get bored quickly and lose interest.	1	2	3	4
32.	I can get over things that would traumatize others.	1	2	3	4
33.	I am sensitive to the feelings of others.	1	2	3	4
34.	I have conned people to get money from them.	1	2	3	4
35.	It worries me to go into an unfamiliar situation without knowing all the details.	1	2	3	4
36.	I don't have much sympathy for people.	1	2	3	4
37.	I get in trouble for not considering the consequences of my actions.	1	2	3	4
38.	I can convince people to do what I want.	1	2	3	4
39.	For me, honesty really is the best policy.	1	2	3	4

40.	I've injured people to see them in pain.	1	2	3	4
41.	I don't like to take the lead in groups.	1	2	3	4
42.	I sometimes insult people on purpose to get a reaction from them.	1	2	3	4
43.	I have taken items from a store without paying for them.	1	2	3	4
44.	It's easy to embarrass me.	1	2	3	4
45.	Things are more fun if a little danger is involved.	1	2	3	4
46.	I have a hard time waiting patiently for things I want.	1	2	3	4
47.	I stay away from physical danger as much as I can.	1	2	3	4
48.	I don't care much if what I do hurts others.	1	2	3	4
49.	I have lost a friend because of irresponsible things I've done.	1	2	3	4
50.	I don't stack up well against most others.	1	2	3	4
51.	Others have told me they are concerned about my lack of self-control.	1	2	3	4
52.	It's easy for me to relate to other people's emotions.	1	2	3	4
53.	I have robbed someone.	1	2	3	4
54.	I never worry about making a fool of myself with others.	1	2	3	4
55.	It doesn't bother me when people around me are hurting.	1	2	3	4
56.	I have had problems at work because I was irresponsible.	1	2	3	4
57.	I'm not very good at influencing people.	1	2	3	4
58.	I have stolen something out of a vehicle.	1	2	3	4

Appendix H

Informed Consent Form

The purpose of this informed consent is to ensure that you now understand the true purpose of the study and the true nature of your involvement. This informed consent provides sufficient information for you such that you have the opportunity to determine whether you wish to continue participating in the study. This study has been approved by the Carleton University Ethics Committee for Psychological Research.

Present study: Spot the Violent Sex Offender: Identifying Violent and Less-violent Sex Offenders at a Glance

Research personnel: The principal researcher for this study is Mari Shanahan Somerville (MariShanahanSomervil@email.carleton.ca). The faculty supervisor for this study is Dr. Adelle Forth (telephone: 613-520-2600 ext. 1267. adelle.forth@carleton.ca)

Purpose: This study aims to identify which personality traits are associated with higher accuracy in perceiving violence from just a short glance at a facial photograph. We are also interested in what facial cues individuals attend to in order to inform this judgment.

Task requirements: You will be asked to rate offender facial images from 1 (less likely to be violent) to 4 (more likely to be violent) from two second exposures to the images. You will also be asked to complete personality questionnaires in which you will be asked to answer questions related to your own personality traits (e.g., psychopathic traits, dominance, extraversion).

Remuneration: You will receive a 0.50% course credit for your participation in this study.

Time Required: To complete all judgments and questionnaires on-line should take approximately 45 minutes. To read and accept the informed consent and then the debriefing afterwards should take an additional 5 minutes for a total of 50 minutes.

Right to Withdraw: I understand that I have the right to withdraw from this study at any point, without penalty. If I choose to withdraw before the end of the survey, I simply need to click on the 'withdraw' button and it will take me to the debriefing form. I will still receive my 0.25% toward my psychology course.

Potential Risk and Discomfort: I understand that I may feel some stress, anxiety or discomfort when assessing sex offenders' propensities toward violence. I understand that I may choose not to answer any questions which cause stress, anxiety or discomfort, without penalty. I understand that if at any time this distress becomes too great, I may withdraw from the study without penalty.

Anonymity/Confidentiality: All responses will be strictly confidential. I understand that the information will be used for research purposes, including further publications, conference

presentations or teaching material. All responses will be coded in such a way that participants cannot be identified.

We will collect data through the software Qualtrics, which uses servers with multiple layers of security to protect the privacy of the data (e.g., encrypted websites and pass-word protected storage). Please note that Qualtrics is hosted by a server located in the USA. The United States Patriot Act permits U.S. law enforcement officials, for the purpose of an anti-terrorism investigation, to seek a court order that allows access to the personal records of any person without that person's knowledge. In view of this we cannot absolutely guarantee the full confidentiality and anonymity of your data. With your consent to participate in this study you acknowledge this.

This study has received clearance by the Carleton University Research Ethics Board-B (Ethics Approval # 105574)

Concerns: Should you have any ethical concerns about this research, please contact Dr. Shelley Brown (Chair, Carleton University Research Ethics Board – B, shelley.brown@carleton.ca, 613-520-2600, ext. 1505). For additional concerns, please contact the Carleton university research ethics board at ethic@carleton.ca

Consent. I have read the above form and hereby consent to continue participating in this study. The data in this study will be used for research publications and/or teaching purposes. I am aware that the data collected in this study will be kept strictly confidential and anonymous. By clicking on the box I consent to participate in this study.

- ☐ By checking this box I indicate that I have read the above informed consent and that I am fully aware of any possible risks or consequences of participating in this study.
- ☐ By checking this box I indicate that I do not consent to the study and wish to exit the survey.

Appendix I

Debriefing Form

What are we trying to learn in this research?

This study is intended to expand on the existing thin slice literature. Thin slice research involves examining observers' accuracy in predicting certain outcomes or forming perceptions of others' personality traits (Fowler, Lilienfeld, & Patrick, 2009). Much of this research is conducted based on the premise that certain personality characteristics may lead to more accurate detection of personality traits from just a short exposure to a face, interaction, or behaviour (e.g., Short, Mondloch, McCormick, Carré, Ma, Fu, & Lee, 2012).

We used measures of the personality dimensions of honesty-humility, extraversion, agreeableness, conscientiousness, openness, boldness, meanness, disinhibition, and psychopathy. We will be using your self-reported personality characteristics to run analyses that will allow us to determine which personality characteristics are associated with a greater accuracy in determining violence from short exposures to facial photographs. We will also be using research assistants' ratings of the offenders' facial characteristics to run analyses that will allow us to determine facial features that are more highly associated with a greater propensity toward violence as well as those that are associated with higher ratings by participants on inclination toward violence.

Why is this important to psychologists or the general public?

This research is important because the results will provide novel information in the domain of thin-slice literature. The majority of thin-slice research thus far has centred its focus on what features about the target (in this case, the sex offenders) are associated with evaluations of higher inclination toward violence by participants (raters) as well as those associated with a history of violence. The present study, however, will also place a large focus on the personality characteristics of the raters that are associated with a greater accuracy in assessing others' propensity for violence from facial characteristics. Such a correlation could have implications for research on evolutionary adaptations for the avoidance of danger and will expand knowledge in the domain of thin slice research.

What are our hypothesis and predictions?

We expect that facial width to height ratio (FWHR) and features associated with masculinity will be correlated with assessments of high inclination toward violence by raters. We also expect that, consistent with Stillman, Maner, and Baumeister (2009), having a heavier brow, being younger, looking more masculine overall, and exhibiting high physical strength will be associated with greater accuracy of participants' assessments of violence. We predict that appearance of physical strength will be more strongly correlated with expectations of violence than actual history of violence. Additionally, we expect that happy, well-groomed offenders will erroneously be rated as having a lower propensity for violence and that offenders displaying expressions of anger or disgust will erroneously be rated as having a higher propensity toward violence.

Further, we expect that individuals high on psychopathic personality characteristics will be more accurate in their assessments of violence. Analyses conducted on the personality domains of honesty-humility, extraversion, agreeableness, conscientiousness, and openness will be exploratory.

Where can I learn more?

For more information on thin slice research, you can read the following studies:

Fowler, K.A., Lilienfeld, S.O., & Patrick, C. J. (2009). Detecting psychopathy from thin slices of behavior. *Psychological Assessment*, 21, 68-78. doi: 10.1037/a0014938

Stillman, T. F., Maner, J. K., Baumeister, R. F. (2009). A thin slice of violence: Distinguishing violent from nonviolent sex offenders at a glance. *Evolution and Human Behaviour*, 31, 298-303. doi: 10.1016/j.evolhumbehav.2009.12.001

What if I have questions later?

If you wish to discuss this research any further feel free to contact Mari Shanahan Somerville (Principal Investigator, email: MariShanahanSomervil@cmail.carleton.ca) or Dr. Adelle Forth (Faculty sponsor, 613-520-2600, ext. 1267; email: adelle.forth@carleton.ca). If you should have any ethical concerns about this study please contact, Dr. Shelley Brown (Chair, Carleton University Research Ethics Board – B, 613-520-2600, ext. 1505; email: shelley.brown@carleton.ca). Should you have any other concerns please contact Carleton University Research Ethics Board – B at ethics@carleton.ca

This study has received clearance by the Carleton University Research Ethics Board-B (Ethics Approval 105574). Please use this number if you need to contact the Chair of Ethics Committee concerning this study.

Is there anything that I can do if I found this study to be emotionally draining?

If you find that after participating in this study you are feeling any sort of emotional, mental or physical stress or anxiety please contact the Carleton University Health and Counseling Services at 613-520-6674. If you are in the Ottawa area, you can also contact Distress Center of Ottawa and Region at 613-238-3311. If you are located somewhere else, the telephone number for Crisis Lines can be found at the front of most phone books.

Thank you for taking the time to take part in this study.

Your participation has been greatly appreciated.

Appendix J**Propensity for Violence Likert Scale**

Please rate the propensity for violence of the individual in the photograph you just saw.

- 1) Unlikely to be violent
- 2) Somewhat unlikely to be violent
- 3) Somewhat likely to be violent
- 4) Likely to be violent

Appendix K

Correlations of Cues with PLV by Gender and Actual Target Violence

Table K.1

Correlations of Cues with PLV by Gender and Actual Target Violence (Sample 1)

Cue	Female PLV	Male PLV	Actual Violence
FWHR	.12	.08	-.14
Smile	.14	.07	-.01
Absence of glasses	.18	.29**	.22*
Jewelry	-.01	-.09	-.14
Facial hair	-.11	-.14	-.03
Long hair	.10	.11	.10
Attractive	.05	.19	.37**
Creepiness	.27*	.08	-.27*
General facial masculinity	.19	.30**	.28*
Heaviness of brow	.33**	.37**	.31**
Care in grooming	-.13	-.03	.38**
Appearance of physical strength	.32**	.42**	.11
Quality of photo	-.28*	-.25*	-.15
Appearance of physical handicap	-.16	-.17	-.25*
Appearance of mental disorder	-.02	-.02	-.32**
Facial symmetry	-.21	-.17	.31**
Older Age	-.31**	-.42**	-.26*
Appearance of arrogance	.25*	.28*	.10
Angry expression	.44**	.33**	-.12
Fear expression	-.15	-.15	.11
Happy expression	-.15	-.07	.03
Sad expression	-.02	-.02	.15
Surprise expression	-.18	-.19	-.05

Note. PLV = Perceived likelihood of violence.* $p < .05$. ** $p < .01$.

Table K.2*Correlations of Cues with PLV by Gender and Actual Target Violence (Sample 2)*

Cue	Female PLV	Male PLV	Actual Violence
FWHR	.08	.05	-.14
Smile	.18	.13	-.01
Absence of glasses	.15	.15	.22*
Jewelry	.01	-.06	-.14
Facial hair	-.03	-.11	-.03
Long hair	.11	.20	.10
Attractive	.09	-.03	.37**
Creepiness	.21	.31**	-.27*
General facial masculinity	.20	.13	.28*
Heaviness of brow	.36**	.33**	.31**
Care in grooming	-.10	-.15	.38**
Appearance of physical strength	.33**	.21	.11
Quality of photo	-.30**	-.32**	-.15
Appearance of physical handicap	-.22*	-.10	-.25*
Appearance of mental disorder	-.05	.08	-.32**
Facial symmetry	-.16	-.25	.31**
Older age	-.33**	-.25*	-.26*
Appearance of arrogance	.26*	.22*	.10
Angry expression	.46**	.42**	-.12
Fear expression	-.10	-.10	.11
Happy expression	-.18	-.15	.03
Sad expression	-.02	.03	.15
Surprise expression	-.20	-.13	-.05

Note. PLV = Perceived likelihood of violence.* $p < .05$. ** $p < .01$.