A LIFE HISTORY MODEL OF HUMAN FITNESS INDICATORS

by

Jon Adam Sefcek

Copyright © Jon Adam Sefcek 2007

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PSYCHOLOGY

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

WITH A MAJOR IN ETHOLOGY AND EVOLUTIONARY PSYCHOLOGY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2007

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Jon A. Sefcek entitled A Life-History Model of Human Fitness Indicators and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

James E. King Date: 8/13/07

Aurelio Jose Figueredo Date: 8/13/07

W. Jake Jacobs Date: 8/13/07

H. Dieter Steklis Date: 8/13/07

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College. I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director: James E. King Date: 8/13/07

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Jon Adam Sefcek

ACKNOWLEDGEMENTS

I would like to begin by thanking my committee for their guidance and support not only on this project but also for preparing me for a scientifically-minded career in psychology. Specifically, I would like to thank Dr. James King for his primatological expertise and support with my various chimpanzee projects; Dr. A.J. Figueredo for his vast statistical and design expertise; Dr. Jake Jacobs for hours of practical and theoretical discussion as well as career advice over various delicious libations; and Dr. Dieter Steklis for his interest and enthusiasm in joining my committee at such a late date.

I would also like to thank Dr. Geoffrey Miller for his theoretical guidance, humorous discussions, and help in fashioning a teaching style that is informative and entertaining for students; Dr. Virginia Landau for her selfless aid in helping me to collect data for various research projects involving chimpanzees, and Dr. Stephen Zegura for his encyclopedic knowledge of evolutionary science and its application to all aspects of scientific inquiry has not only helped buttress my breadth of interests, but also given me a safe haven outside the halls of psychology.

Finally, I would like to thank my friend Pedro Wolf who has not only helped me struggle through theoretical questions and data analysis, but also helped keep me grounded in the real world outside the basement of our little ivory tower; My brother, Rob, who's refusal to call me Master, has driven me to become a Doctor and Professor; my mother Sally and father Bob who have always been supportive of my academic pursuits; and my friend Jill Endres, who has taught me that there are some things important enough to pursue to fruition, despite how difficult they are to accomplish.

DEDICATION

This dissertation is dedicated to memory of my father Robert W. Sefcek,
who despite his low- K heritage,
gave me good genes,
a supportive family,
and,
unwavering parental investment.

TABLE OF CONTENTS

LIST OF TABLES.....	8
LIST OF FIGURES	9
ABSTRACT.....	10
INTRODUCTION AND LITERATURE REVIEW	11
Overview.....	11
General Intelligence as a Latent Construct.....	12
Life History Theory: an r/K perspective.....	13
Pleiotropic Mutation and Fitness Indicator Theory	18
An Integrated Evolutionary Model.....	21
Specific Predictions	22
<i>Predictions deriving from Differential-K Theory</i>	22
<i>Predictions deriving from the General Fitness Model</i>	23
<i>Predictions derived from the Hybrid Model</i>	23
Explanation of Dissertation Format.....	23
PRESENT STUDY	26
Information from the Pilot Study.....	26
Method	27
<i>Participants</i>	27
<i>Scales and Measures</i>	27
<i>Measures of Somatic Investment</i>	27
<i>Measures of Morpho-developmental stability</i>	29
<i>Measures of Neuro-developmental Stability (Mental Health)</i>	32
<i>Measures of Neuro-developmental Stability (Intelligence)</i>	34
Statistical Analyses	36
Results.....	36
<i>Scale Reliabilities</i>	36
<i>Common Factor Models</i>	37
<i>Exploratory K-Factor Loadings</i>	37
<i>Exploratory F-Factor Loadings</i>	38
<i>Exploratory g-factor Loadings</i>	39
<i>Unit-weighted Higher Order Factors</i>	39
<i>Multiple Regression Models</i>	40
<i>Tests of the Differential-K Model</i>	41
<i>Tests of the General Fitness Model</i>	42
<i>Test of the Hybrid Model</i>	42
Discussion.....	43
Outcomes of Specific Predictions.....	43
<i>Prediction 1</i>	43
<i>Predictions 2 & 4</i>	46
<i>Predictions 3, 5, & 6</i>	48
Limitations and Future Directions	52
REFERENCES	54

TABLE OF CONTENTS CONTINUED

TABLES	64
FIGURES	74
APPENDIX A: CHIMPANZEE FACIAL SYMMETRY: A BIOMETRIC MEASURE OF CHIMPANZEE HEALTH.....	79
Abstract	79
Introduction.....	80
Method	82
Results.....	85
Discussion	86
Acknowledgements.....	88
References	88
Table 1	91
Figure 1	92
APPENDIX B: DEVELOPMENT AND VALIDATION OF AN 18-ITEM MEDIUM FORM OF THE RAVENS ADVANCED PROGRESSIVE MATRICES (APM-18)	93
Abstract	93
Introduction.....	94
Study 1: Scale Construction and Construct Validity	95
Method	95
Results.....	97
Discussion	98
Study 2: Convergent Validity	99
Method	99
Results.....	103
Discussion	104
Conclusion	105
References.....	105
Tables.....	108

LIST OF TABLES

Table 1. Descriptors of traits associated with <i>r-selection</i> and <i>K-selection</i>	64
Table 2. Sample size comparisons and Cronbach's alpha's across measures.....	65
Table 3a. Factor loadings for two-factor solution of <i>K-Factor</i>	66
Table 3b. Factor loadings for single-factor solution of <i>F-Factor</i>	67
Table 3c. Factor loadings for single-factor solution of <i>g-factor</i>	68
Table 4. Higher-order factor loadings (reported as standardized regression coefficients).....	69
Table 5a. Correlations among higher-order factors	70
Table 5b. Correlations among higher order factors, by sex	71
Table 6. Correlations with particular 'g' measures.....	72
Table 7. Multiple Regression Statistics for the Theoretical Models.....	73

LIST OF FIGURES

Figure 1. Example of the interaction between the G-matrix and M-matrix.....	74
Figure 2. Differential-K model and general intelligence	75
Figure 3. Fitness factor model and general intelligence	76
Figure 4. Hybrid model of the relationships among K , F , and g	77
Figure 5. Points used in calculating facial asymmetry measurements.....	78

ABSTRACT

Recent adaptationist accounts of human mental and physical health have reinvigorated the debate over the evolution of human intelligence. In the tradition of strong inference the current study was developed to determine which hypothesis; Rushton's (2000) differential K theory, or Miller's (2000a) fitness indicator model (F), better accounts for general intelligence (' g ') in an undergraduate university population ($N = 194$). Due to the lengthy administration time of the test materials a newly developed 18-item short form of the Ravens Advanced Progressive Matrices (APM-18; Sefcek, Miller, & Figueredo, 2007) was used. There was a significant positive relationship between K and F ($r = .31, p < .001$), however no significant relationships were found between ' g ' and either K or F (for each, $r = -.06, p \geq .05$). While contrary to both hypotheses, these results may be explained in relation to antagonistic pleiotropy and a potential failure to derive within species comparisons directly from between species comparisons.

INTRODUCTION AND LITERATURE REVIEW

Overview

Few topics have stirred such controversy in both academia and the popular press as the study of individual variation in human intelligence. This has been especially true when such differences are documented based on biological and genetic mechanisms (Segerstråle, 2000). Indeed, the polemic debates that arise from such scientific pursuits have often led to cries of ethnocentrism, racism, sexism; which can lead to the renouncement of theories and destruction of careers (Gould, 1981). However, recent advances in behavioral and genetical research have yielded strong evidence that variation in intelligence is heritable and that a scientific understanding of this variation must take into account evolutionary arguments (Plomin, 2003).

Applying this perspective to human intelligence, many evolutionary accounts for both its emergence and within species variation have arisen (see Bock, Goode, & Webb, 2000, for several examples). This paper will focus on two of the most prominent of these. The first draws upon the survival-side of natural selection by exploring the selective pressures (e.g. changing climactic conditions, predictable seasonality) that necessitated the development of adaptable, large-brained, intelligent humans. The second is a more direct focus on the reproductive benefits of human intelligence, shaped through sexual selection, to serve as an indicator of genetic quality. The purpose of the current research is to apply strong inference (Platt, 1964) and the method of multiple working hypotheses (Chamberlain, 1965), in a theoretical ‘battle royale’ to determine which best accounts for general intelligence, or ‘g’.

General Intelligence (‘g’) as a Latent Construct

Beginning with Galton (1869, 1962) there has been a consensus among intelligence researchers and psychometricians that there is a *general* cognitive ability, important to many pursuits in life inclusive of academic and work achievements as well as developmental (life-history) trajectories. This general dimension was empirically validated through factor analytic techniques by Spearman (1904), who identified high positive intercorrelations among many different facets of intelligence. Since its inception, this general intelligence, or ‘g’, has been shown to relate to a variety of psychosocial (e.g. age of school dropout, poverty, unplanned pregnancy, criminality, health risk behavior), and biological/information processing traits (e.g. mass of prefrontal cortex and overall brain mass, nerve conductivity velocity, reaction time, inspection time, and glucose metabolism rates, health outcomes, longevity), in the predicted manner, even when SES is controlled in the research design (Murray, 1998).

However, it is generally accepted among intelligence researchers that ‘g’ is not one specific trait, or module, that resides somewhere in the brain (see Kanazawa, 2003, for alternate viewpoint). Rather, it is a latent variable that is comprised of some ability that allows an organism to tap into, and integrate, various aspects of cognitive function (e.g. learning, memory, language, reasoning; Jensen, 2000). No one can see ‘g’ specifically; however the phenotypic correlates of ‘g’ can be seen. Variability arises because some individuals have the inherited capacity to integrate these various areas more quickly and with more efficiency than others. Theoretical considerations as to the etiology of this ability are as varied as the ability itself. However, the most prominent of

intelligence researchers do agree that this ability is the result of evolutionary pressures (Bock, et al. 2000).

Life-History Theory: an r/K perspective

Life History Theory is a mid-level evolutionary theory which examines the manner in which organisms allocate material (e.g. calories, time) and bioenergetic (e.g. energy) resources among the various demands of survival and reproduction (Figueredo, Vásquez, Brumbach, Schneider, Sefcek, Tal, et al., 2006). These allocations may take the form of *somatic effort*, i.e. investing in personal survival and longevity, and *reproductive effort*, i.e. investing in mating, parental, and kin investment (McArthur & Wilson, 1967). At its core it is an economic theory that describes how organisms strategically allocate limited resources among the survival of currently existing genes (self, offspring, and genetic relatives) and the potential production of new genes (mating) (MacArthur & Wilson).

One interpretation of Life History Theory has been explained through r/K selection. In its purest form r/K theory describes the tradeoff between two ends of a continuum where parental investment is high and fecundity low (K -selected), and fecundity is high and there is no parental investment (r -selected; Figueredo, Vásquez, Brumbach, Sefcek, Kirsner, & Jacob, 2005). Associated with these varying patterns of investment/fecundity are a variety of physical and behavioral traits that are argued to aid long term survival (K -selected) or short-term reproduction (r -selected; *Pianka*, 1970; see Table 1, for r versus K descriptors). According to Pianka, the selective pressures shaping these different patterns are the stability and predictability of the environment.

For example, organisms which evolved in unstable, unpredictable environments would have evolved clusters of traits (e.g. high fecundity, low parental investment, low inter-birth interval, rapid sexual maturity) that produced a high quantity of offspring set to reproduce as quickly as possible in such an uncertain environment. On the contrary, organisms which evolved in stable, predictable environments would evolve clusters of traits (e.g. low fecundity, high parental investment, high inter-birth interval, delayed sexual maturity) that produce a few high-quality offspring in a relatively worry-free environment. Due to an evolutionary legacy within these varying environmental niche, traits associated with increased fitness, for that niche) will have become phenotypically and genetically correlated to each other in an overarching, observable, behavioral pattern (Figueredo, Vásquez, Brumbach, & Schneider, 2004).

Traditionally this concept has been applied to variation in behavioral patterns between-species rather than within-species. However, the theory itself predicts that organisms would have the ability to track their ecological condition and to alter their strategies in ontogenetic time. In light of such a prediction, Rushton (1985) has examined this proposition through what has been termed *Differential-K theory*. This theory posits that that ability to allocate ones resources differentially in response to environmental condition is an adaptation allowing individuals to track and respond to environmental conditions, which will express itself phenotypically as individual within-species variation. Indeed, in a review of key life-history traits Rushton has shown that this sort of within-species variation does exist in diverse species such as; dandelions (Gadgil & Solbrig, 1972), field mice (Krebs, Gaines, Keller, Myers, & Tamarin, 1973), American

shad (Leggett & Carscadden, 1978), as well as flies and milkweed bugs (Taylor & Condra, 1980).

More recently, *Differential-K theory* has been applied to the study of variation in human life-history traits (called the *K-Factor*) in a life-history focused research program at the University of Arizona (Figueredo, Sefcek, Vasquez, Brumbach, King, & Jacobs, 2005; Figueredo, Vásquez, Brumbach, Sefcek, et al., 2005; Figueredo, Vásquez, Brumbach, Schneider, et al., 2006; but see Figueredo, Vásquez, Brumbach, Schneider, Sefcek, et al., 2006, for review). This program owes its original theoretical basis to Rushton (2000) where it was argued that environmental shifts associated with human migration out of Africa some 60-80,000 years ago (Oppenheimer, 2003) led to harsher environments that necessitated the development of higher cognitive abilities (e.g. intelligence, executive function). As climates became increasingly cooler and seasonal, new means of securing food and shelter became necessary.

Such harsh ecological factors led not only for a need for higher intelligence, but also higher male parental investment to care for altricial young, stronger group cohesion, and enhanced forethought and planning ability to predict seasonal, yet stable, ecological patterns. Therefore, the expanse of cognitive abilities necessary to accomplish such feats led to a high-*K* lifestyle typified by a slow reproductive strategy and high parental and kin investment. This model also makes the counter prediction that individuals whose ancestors evolved in unstable, unpredictable environments would tend to invest their metabolic resources into early, frequent reproduction at a cost to a well-built, long-lived body, leading to a low-*K* lifestyle (Rushton, 2000, 2004).

Additionally, psychometric traits, e.g. intelligence, personality, psychopathology, and many physiological markers of health and immune function (e.g. negative health symptoms, general physical health, facial and body fluctuating asymmetry), should have been shaped by natural selection to produce meaningful functional composites. If evolution did shape a subset of individuals to develop the cognitive architecture to focus on long-term reproductive strategies for example, then it should have shaped bodies that are well-prepared to survive long-term (high intelligence, strong immune systems, lower indices of negative health), rather than having a motley assortment of traits that may work in an antagonistic fashion. Indeed, recent research has shown strong positive phenotypic and genetic relationships between the *K-Factor*, and a global health factor (denoted ‘Covitality’; Figueredo et al., 2004; Figueredo, Vásquez, Brumbach, & Schneider, 2007; Figueredo, Vásquez, Brumbach, Sefcek, et al., 2005)

Many traits associated with *Differential-K theory* are heritable in humans. For example, fecundity (Rodgers, Hughes, Kohler, Christensen, Doughty, & Rowe, 2001), pubertal timing and menarchal age (Rowe, 2000), and age of first reproduction and menopause (Kirk, Blomberg, Duffy, Heath, Owens, & Martin, 2001), have heritabilities ranging between (.23 - .50). More recently, work examining the *K-Factor* has reported a factor heritability of .65 (Figueredo, et al., 2004).

While genetic contributions to reproductive strategy are incontrovertible, the associated environmentality is still high enough to allow for quite a degree of plasticity during ontogeny (Figueredo, Vásquez, Brumbach, Schneider, et al., 2006). Indeed in several studies (e.g. Belsky, Steinberg, & Draper, 1991; Chisholm, 1996; Ellis, 2004)

environmental factors such as father absence or stepfather presence has been shown to relate to a low-*K* lifestyle indicated by earlier menarche, earlier age of sexual intercourse and reproduction, and shorter-term and unstable romantic relationships. This is consistent with the predictions derived from *Differential-K theory*.

Other, yet to be explored predictions stemming from this theory, include the relationship between the *K-Factor* and behaviors associated with psychopathy and social deviance. For example, many researchers suggest that the behaviors associated with psychopathy (e.g. Machiavellianism, impulsivity, ignorance of the costs of risky behaviors, lack of social emotions, hostile aggressiveness) may work as an intraspecific cheater phenotype focused on short-term, opportunistic mating (e.g. Mealey, 1995; Weibe, 2004). As Hare (1993) points out, real psychopaths tend to be promiscuous, philandering, and have multiple children with multiple partners. These are just the types of behaviors we would expect from individuals adopting an extreme low-*K* strategy.

Literature on psychosocial behaviors and social deviance also supports a relationship to *Differential-K theory* (Ellis, 1988; but see Figueredo, Vásquez, Brumbach, Schneider, et al., 2006, for review). A variety of research has shown that criminal and delinquent behaviors (e.g. substance abuse, gang involvement, risky sexual behavior, violence, teen pregnancy, sexual and physical abuse, poor parental supervision) tend to cluster together in an observable behavioral pattern with genetic covariance (*'d'*; Rowe & Rodgers, 1989; *delinquency proness*; Rowe & Flannery, 1994). Rowe, Vazsonyi, and Figueredo (1997) later reported the possibility that heritable differences in life-history

strategy accounted for significant correlations between juvenile delinquency and life-history traits within individuals and across siblings.

Pleiotropic Mutation and Fitness Indicator Theory

An alternate view for the origin of individual differences in cognitive capacities (e.g. intelligence) is the general fitness model (F), proposed by Miller (1998, 2000a, 2000b). This model posits that many psychometric, psychosocial, and physical markers of health are linked together in an overarching fitness factor (F -Factor) that directly taps into an individual's underlying genetic quality, or mutation load. Each of these variables is part of one of four larger constructs in which the individual may allocate its bioenergetic and material resources: 1) somatic investment, the way in which you invest resources in body and brain growth; 2) morpho-developmental stability, or how resistant the genes instructing your bodily development are to the onslaught of negative stressors, e.g. parasites, toxins, malnutrition, and mutation; 3) neuro-developmental stability, how strong the genes instructing your cognitive development are at resisting negative stressors; and 4) Immunocompetence, or how well your immune system, as a whole, fends-off disease. Each of these constructs is assumed to show some degree of developmental stability, i.e. reflect resistance to environmental variation and perturbations, because regardless of environment, those with fewer deleterious genetic mutations will outperform those with more mutations. Therefore, life-history strategy is not the cause of intelligence or these other factors but may simply be another manifestation of general fitness.

This model utilizes pleiotropic mutation theory, a quantitative genetics model that

describes the manner in which genes comprising the current phenotypic/genetic variance-covariance within a population (G-matrix) and the potential for change to this matrix through a few mutational steps (M-matrix) work antagonistically with each other (Houle, 1991; Pigliucci & Hayden, 2001). Figure 1 represents a simplified example of this model. In this figure each of the eight alleles (a1 – a8) affects at least one-trait (T1 – T4) in an additive way, but several alleles affect more than one trait (e.g. a1, a2, a4) producing a pleiotropic effect on these traits. The traits therefore become linked together due to this web of allelic influence. If a mutation arises on only 2-alleles out of 8 (indicated by the dashed-lines of a1 and a4) we see that in fact all four traits are affected by the mutations, with T1, T2, and T4 being affected more strongly due to the number of mutations affecting them.

As Rowe & Houle (1996) point out, the larger the *mutational target size* (the more genes that are associated with any one trait) then the better chance there is that any deleterious mutation will be shown phenotypically. This means that those traits associated with fitness, through pleiotropy, will show lower phenotypic quality than the population average when an individual has a higher than average number of mutations acting on the alleles coding for the traits. Further, it suggests that there should be more phenotypic variance on traits among low-fitness individuals than high fitness ones.

This is essentially the genetic basis for fitness indicator theory (Miller, 1999). Fitness indicator theory posits that many phenotypic traits used in mate choice, friend choice, competition for access to mates, and predator avoidance, serve as honest signals of genetic quality as measured by deleterious mutation load. These indicators must

accurately and reliably express condition to others, operating under Zahavi's (1975, 1991) handicap principle. For example, stotting in gazelles signals to potential predators "Hey I have abundant energy you will never catch me", and to females "I have so much energy that I can afford to waste it on this behavior" (Miller, 2000). However, if it is faked the gazelle will be found out by becoming dinner for a cheetah after it has wasted all its energy in a false display. This is why fitness indicators do the voodoo they do so well. They are conspicuously wasteful to produce in terms of metabolic energy, time, and risk; and because they tap numerous alleles the pleiotropic affect of the alleles, and any deleterious mutation, on said indicators will be intimately linked. And because they are condition dependant (responsive to environmental stressors), the most informative indicators will show the highest phenotypic variation in size, complexity, or quality across individuals within the same environment. Due to the fitness enhancing effects of these indicators, they will, over evolutionary time, become genetically and phenotypically correlated to each other.

In Miller's (2000) view the connection between pleiotropic mutation theory and fitness indicator theory is why the human brain, and facets of it, e.g. creativity, intelligence, psychopathology, and personality, has evolved. Due to its large mutational target size (approximately 50% of human genes affect brain architecture in some way), it is argued that the likelihood that some genetic mutations will be shown through the products of the brain, i.e. behavior, is high (Rowe & Houle, 1960). Coupled with its small size (2% of body weight) versus high metabolic cost (requiring approximately 15% of oxygen, 25% of calories, and 40% of glucose), and ease of disruption due to poor

nutrition, disease, injury, inbreeding, and social isolation (Miller), it is arguably the most interesting and elaborate fitness indicator known to science.

Further, because of such a large mutational target size, it is likely that many of these genes have a pleiotropic effect on other aspects of an organism (e.g. immune function and health), suggesting observable correlations among these various traits. If we go back to Figure 1, for example, and assume that T1 – T4 were different aspects of cognitive function (e.g. creativity, intelligence, personality, psychopathology), that were coded for by different alleles forming brain architecture (a1 – a8) then we can begin to see how genetic mutations of underlying brain architecture may affect cognitive function, and how such mutations may be phenotypically expressed in a variety of traits. Therefore, we should see both genetic and phenotypic correlations between indicators of fitness. These things may range from immune function and markers of health, to personality, psychopathology, and intelligence.

An Integrated Evolutionary Model

Currently, which model, i.e. *Differential-K* or general fitness (F), best accounts for phenotypic variation in human mental, physical, and health traits is uncertain. Several lines of evidence support each. Both theories suggest that the use of multivariate correlational techniques will be able to detect underlying functional latent constructs and behavioral signatures. *Differential-K* theory predicts that each of the constructs outlined above; somatic investment, morpho-developmental stability, neuro-developmental stability, and immunocompetence should be positively related producing individuals who are high- K , healthy, symmetrical, and cognitively adept (see Figure 2). The General

Fitness model (F) predicts that underlying genetic fitness should be driving somatic investment, morpho-developmental stability, neuro-developmental stability, and immunocompetence, therefore high-quality indicators should all be positively related. It also suggests that those with the best genetic fitness will produce healthier, higher-fitness offspring, who themselves may enjoy reproductive advantages (see Figure 3).

One of the models may best account for individual differences in variation on these traits. However, the simple fact is that these models are not mutually exclusive (see Figure 4). Rather they examine human traits from two different perspectives (e.g. survival and parental investment versus sexual selection), both of which have strong evolutionary effects.

The current study was developed to examine if either model better accounts for variation in human mental capacities, or if the two of them are themselves related. This will be examined partially by extending research that established the validity of measuring life history strategy via self-report measures of family configuration and support, kin and friend altruism, adult romantic attachment, reproduction, mating effort, and risk taking (Figueredo, Sefcek, et al. 2005; Figueredo, Vasquez, Brumbach, et al. 2006).

Specific Predictions

Predictions deriving from Differential-K Theory

Prediction 1. Social Deviance and Psychopathic attitudes should load negatively on the *K-Factor*.

Prediction 2. The *K-Factor* should be positively associated and predictive of the

F-Factor (inclusive of good physical health, good mental health, and low fluctuating asymmetry).

Prediction 3. The *K-Factor* should be predictive and positively associated with general intelligence, as measured by the *g-factor*.

Predictions deriving from the General Fitness Model

Prediction 4. The *F-Factor* (including good physical health, good mental health, and low fluctuating asymmetry) should be predictive and positively associated with the *K-Factor*.

Prediction 5. The *F-Factor* should be predictive and positively associated with the *g-factor*.

Predictions derived from the Hybrid Model

Prediction 6. Based on pleiotropic mutation theory the *K-Factor* and the *F-Factor* should be related to each other and should jointly predict general intelligence.

Explanation of Dissertation Format

This dissertation will be comprised of three main sections (present study, Appendix A, and Appendix B). The present study will build off of the hybrid model outlined above. It will explore the predictions of the competing hypotheses, via a research project that was conducted at the University of Arizona's Main Campus in Tucson. This section will begin with a brief explanation of the information collected via a pilot study that was conducted prior to the final sample being collected. This pilot study was important to the present study because the final methodology and analyses were determined by information gathered.

The second and third sections (i.e. Appendix A and Appendix B) are reproduced copies of original research articles (the first one published, the second one submitted for review) in which I am the primary researcher and author. Each of these articles reports on original research projects that were conducted as side-projects related to the present study. Outlined below is a brief explanation of how each of these papers is related to the theoretical framework weaved in the present study.

Appendix A is comprised of a recently published research manuscript (Sefcek & King, 2006) which develops and examines a methodology for measuring facial fluctuating asymmetry in a population of zoo chimpanzees. If, as predicted by fitness indicator theory, facial fluctuating asymmetry is an indicator of genetic and physical health in humans, it is possible that it should also serve as such an indicator of genetic and physical health in other closely related species (e.g. *Pan Troglodytes*, the Chimpanzee).

A large body of literature has been published on the relationships among various genetic and physical health measures and various measures of fluctuating asymmetry (e.g. wings, digits, canines) in non-human species (see Sefcek, Brumbach, Vasquez, and Miller, 2006, for review). However, none of them have incorporated a measure of fluctuating facial asymmetry in non-humans. This paper adapted a methodology reported by Jacobson, Steklis, Steklis, LeClair, & Fawcett (2004), for measuring facial fluctuating asymmetry in wild gorillas for use as a biomonitoring tool (although they have yet to report any findings in relation to the health of these individuals).

We explored the relationship between these asymmetry measures and measures of

general health and negative health symptoms. Results were as predicted, with a composite of chimpanzee fluctuating facial asymmetry being positively related to negative health symptoms and negatively related to overall indices of good health. As a further connection to the present dissertation study, the health measures utilized in the published journal article were adapted from the same health scales used in the present study. Ultimately, a final comparative analysis will be conducted utilizing those data published in Sefcek & King's (2006) study, and unpublished human facial fluctuating asymmetry data collected as part of the present study.

Appendix C is comprised of a submitted paper (Sefcek, Miller, & Figueredo, 2007), which discusses the creation and validation of a medium-form (18-item) test of general intelligence ('g'). This test was created originally for use in the present study, as an attempt to lessen the already lengthy administration time of the measurement instruments, while retaining an acceptable level internal consistency for the scale.

This APM-18 was adapted from the 40-minute long Raven's Advanced Matrices (APM), which is designed to be a culturally fair test of general intelligence. While there is one short-form (12-item) version of this scale available (Arthur & Day, 1994), we determined the reliability of the scale to be lacking for our current purposes (Cronbach's alpha's range between .58 - .66; as reported by Arthur, Tubre, Paul, & Sanchez-Ku, 1999). As reported in our paper, the newly developed APM-18 has shown adequate reliability in two samples ($\alpha = .73$ and $.79$), convergent validity with other measures of intelligence (e.g. the Mill-Hill Vocabulary Scales, Shipley Institute of Living Scale, GPA), with an average administration time of 17.50 minutes (test limited to 25-minutes).

PRESENT STUDY

Information from the Pilot Study

A pilot sample ($N= 70$) was collected from two University of Arizona branch campus classes. Students participated as part of a classroom demonstration used to anchor several future lectures, including a lecture for debriefing them on the entire theoretical research model. Specific results from this demonstration will not be reported here. However, three important points of information that influenced the present study are outlined below:

- 1) The proposed battery of measurement instruments took more time than originally anticipated (over 2-hours). As such, several people did not complete several portions of the study, which was limited to 2-hours.

- 2) There was little variance associated with the scores on the only pure test of general intelligence being used, the Ravens Standard Progressive Matrices (SPM).

Consequently, the internal consistency of this scale was uninterpretable, leading to a problem with the *g-factor* as a whole. Because of the lengthy administration time and the problems with the SPM, a new mid-length form of the Ravens Advanced Progressive Matrices (APM) was developed (see below for brief explanation, Appendix B for complete paper). The aim of developing this scale was to 1) lessen the administration time necessary for a “pure” measure of ‘g’, and 2) to develop a scale that would be better at measuring individual variation in general intelligence than the SPM, leading to higher internal consistency.

- 3) Relationships previously reported for K and F were not identified. Specifically, there

was no significant correlation found between these two latent factors, although they had been reported in other samples (for example, see Figueredo, et al., 2005). After contemplating the theoretical nature of the attenuated relationship between K and F we decided to conduct an exploratory factor analysis (EFA) for each cluster of scales that were theoretically predicted to comprise each factor.

Method

Participants

One-hundred ninety-four (100 female, 94 male) introductory psychology students at the University of Arizona, Main Campus, participated in this study. All participants were given the option of fulfilling a course requirement through research participation or writing a research-oriented paper. The mean age of participants was 19.11, $SD = 1.62$ ($M_{female} = 19.01$, $SD_{female} = 2.00$; $M_{male} = 19.23$, $SD_{male} = 1.07$). Due to lengthy administration times of all materials, several individuals did not complete one or more measures of the study (see Table 2 for sample size of specific measures).

Scales and Measures

This study utilized a variety of measures devised to tap into the four different hypothesized areas of “general fitness”: 1) Somatic investment, 2) Morpho-developmental Stability, 3) Neuro-developmental Stability and 4) Physical Health (as a proxy for immunocompetence).

Measures of Somatic Investment

Arizona Life History Battery (Figueredo, Sefcek, et al., 2005).

This life-history battery was developed at the University of Arizona to measure

individual differences in life-history strategy. It contains several scales measuring parental attachment and support, support towards and from friends, romantic partner attachment, mating effort, Machiavellianism, and risk taking. To date this battery has indicated strong convergent and discriminant validity, as well as internal consistency (IC) within scales ($\alpha \geq .54$; see Table 2 for IC's for all measures).

While this battery generally contains a Machiavellianism scale, it was substituted with the Psychopathic Personality Inventory (see below) for two reasons. First, one of the main subscales of the battery is specifically related to Machiavellian Egocentricity, allowing for this direct test, if needed. Second, several lines of research based on evolutionary perspectives (e.g. Figueredo, Brumbach, Jones, Sefcek, Vásquez, Jacobs, 2006; Mealey, 1995) suggest that psychopathy may be an adaptive cheater strategy focused on short-term, opportunistic mating. As such, this scale should show a negative relationship to the *K-Factor*.

The Mini-K: Short Form (Figueredo, Vásquez, Brumbach, et al., 2006).

The Mini-K is a 20-item, self-report short form of the *K-battery*. It was developed to be used in situations where time-constraints made use of the complete *K-battery* prohibitive. To date it has been used in a variety of cross-cultural samples with reported Cronbach's alpha and test-retest reliabilities of around .70 (Brumbach, Figueredo, & MacDonald, 2005; Sefcek et al., 2005, Sefcek, Figueredo, & Miller, 2005; Vásquez, Sefcek, Douglas, & Figueredo, 2005; Wenner et al., 2005; Wolf, Vásquez, Frías-Armenta, Corral-Verdugo, & Figueredo, 2005).

The Psychopathic Personality Inventory: Short Form (PPI-SF; Lilienfeld &

Andrews, 1996).

This 56-item scale was derived as a shortened version of the Lilienfeld & Andrews (1996) 187-item self-report questionnaire of non-criminal population psychopathy. This scale was developed to measure personality traits rather than anti-social behavioral propensities. Each item is endorsed on a 1 to 4 scale (1 = False, 2 = Mostly False, 3 = Mostly True, 4 = True), and tallied to produce either a total PPI score, or one of eight subscale-scores: (1) Machiavellian Egocentricity (manipulativeness), (2) Social Potency (charm and influence), (3) Coldheartedness (callousness and guiltlessness), (4) Carefree Nonplanfulness (lack of forethought), (5) Fearlessness (risk taking), (6) Blame Externalization (blaming others for mistakes), (7) Impulsive Nonconformity (a disregard for social norms), and (8) Stress Immunity (lack of anxiety). Global PPI-score as well as the subscales have shown strong internal consistency with Cronbach's alpha ranging from .64 to .93 (Chapman, Gremore, & Farmer, 2003; Edens, Poythress, & Lilienfeld, 1998; Lilienfeld & Andrews; Poythress, Edens, & Lilienfeld, 1998).

Measures of Morpho-developmental Stability

Facial Fluctuating Asymmetry.

Following procedures described by Scheib, Gangestad, & Thornhill (1999) two color digital facial photographs were taken from each participant at a distance of 1-meter, with picture zoom set to 3.0. The degree of asymmetry was then measured via the Image-J 1.31v program of the National Institute of Health (NIH, 2003). First, the midpoints of seven lines (D1 – D7) running horizontally across the face; D1) at the outer and D2) inner

eye corners, D3) the eye pupils, D4) the cheekbones, D5) the outer edges of the nose and D6) mouth, the D7) jaws (at the mouth line) will be calculated (see Figure 5, for example of measurement points). The absolute values of these measurements were then summed to define a horizontal fluctuating asymmetry composite score. Second, the absolute values of the vertical differences were summed to define a vertical fluctuating asymmetry score on all measurement points except the pupils. Because these “horizontal” and “vertical” components are measured on different scales and will have different standard deviations, they will be standardized before summing them into a composite measure of facial fluctuating asymmetry (FFA).

Body Fluctuating Asymmetry.

Body asymmetry measurements are taken with digital calipers reading to the nearest .001 inch, using the methodology described by Thornhill & Gangestad (1999). Participants are seated in a chair with feet flat on the floor while two experimenters measured body symmetry from 10-paired (right and left) locations on body and head. In order to account for any carry-over biases from previous measurement of the paired trait, all measurements are taken down the right-side of the body and then down the left side. These measurements of interest are: 1) ear length, 2) ear width, 3) elbow width, 4) wrist width, 5) digit two length, 6) digit three length, 7) digit four length, 8) digit five length, 9) ankle width, and 10) foot width.

Other studies have reported minor directional biases in measurements. However, only that of the foot has been found to be significant (Gangestad & Thornhill, 2003).

Composite body FA scores are calculated by summing up the relative FA (absolute asymmetry of each paired trait divided by the average trait size) across all traits.

General Health Scale (Brim et al. 2000).

This scale was derived from the MIDUS Life Satisfaction Scale as a global measure of mental and physical health. It is comprised of 5-items rated on either a 0 to 4 (items 1-3) or 0 to 5 (items 4 & 5) scale. Sample items include: “In general, would you say your physical health is...” and “In general, compared to most (men/women) your age, how would you say your health is?” To date the one study using this scale has reported acceptable internal consistency (Cronbach’s $\alpha = .68$; Figueredo, Vásquez, Hagenah-Brumbach, & Schneider, 2005).

Medical Symptoms Scale (Brim et al. 2000).

This scale measures medical symptoms through reports of the endorsement of 28 possible medical conditions. Examples include the endorsement of items including “Thyroid disease,” “High blood pressure or hypertension,” or “Ulcer.” Again the one study which has utilized this scale reports acceptable internal consistency (Cronbach’s $\alpha = .72$; Figueredo, Vásquez, Hagenah-Brumbach, & Schneider, 2005).

The RAND 36-Item Health Survey: Version 1 (SF-36; Ware & Sherbourne, 1992).

This 36-item scale can be used as both a global index of physical and emotional health as well as broken down into eight subscales (based on health concepts): physical functioning, bodily pain, role limitations due to physical health problems, role limitations due to personal or emotional problems, emotional well-being, social functioning, energy/fatigue, and general health perceptions. Subscales may be combined to produce a

physical health summary score (mean of the physical functioning, bodily pain, role limitations due to physical health problems, and general health subscales) and a mental health summary score (role limitations due to personal or emotional problems, emotional well-being, social functioning, energy/fatigue). Typical reports of internal consistency (measured by Cronbach's α) for the eight scales and global index measure have exceeded .80 (McHorney, Ware, Lu, & Sherbourne, 1994; Ware & Sherbourne). Reliability estimates for physical and mental summary scores usually exceed .90. Due to the partial mental health nature of this scale it could also be listed under the heading of

Body Mass Index (BMI; Quetelet, 1842).

BMI, or a person's weight in kilograms (kg) divided by their height in meters (m) squared, is a biometric measure used by health professionals to assess the relative amount of fat versus nonfat composition of the body (Garrow & Webster, 1985). Current National Institute of Health (NIH) guidelines define BMI's as follows: Underweight (below 18.5), Normal (18.5 – 24.9), Overweight (25.0 – 29.9), and Obese (over 30.0) (NIH, 2005). BMI has been found to be highly correlated to numerous chronic health conditions such as cardiovascular disease, high blood pressure, osteoarthritis, various cancers, and diabetes, as well as premature death (Calle, Thun, Petrelli, Rodriguez, & Heath, 1999). As such, it is a good gauge for relative health, although extremely muscular individuals also have high BMI's (minus the negative health risks).

Measures of Neuro-developmental Stability (Mental Health)

Subjective Well-Being Scale (SWB; Figueredo, Vásquez, Hagenah-Brumbach, & Schneider, 2005).

This three-item scale was derived by combining one-item from the MIDUS (National Survey of Midlife Development in the United States) Life Satisfaction Scale (Brim et al. 2000) and two items created specifically to measure general life-happiness. It was developed to assess general SWB, satisfaction with ones life. Internal consistency of this scale is moderate (Cronbach's alpha = .74).

The Beck Depression Inventory: Second Edition (BDI-II; Beck, Steer, & Brown, 1996).

The BDI-II is a 21-item self-report instrument which measures the existence and intensity of depression (as outlined by the DSM-IV) for both general and clinical populations. Nineteen of the items consist of a list of 4-statements arranged in increasing severity (measured 0 – 3) about a corresponding symptom of depression. Two items, 16 and 18, have 7-options associated with increases/decreases in appetite and sleep. Total score of 0-13 is considered minimal range, 14-19 is mild, 20-28 is moderate, and 29-63 is severe, although it is recommended that these categories be adjusted based on the mean of the specific sample. This instrument is designed for use across a variety of age-groups, from 13 – 80. In several studies Cronbach's alpha of the scale is reported to be around .85 (Ambrosini et al. 1991; Craven et al. 1988; Levin et al. 1988; Lovibond & Lovibond, 1995; Naughton & Wiklund, 1993).

The Hopkins Symptom Checklist 25 (HSCL-25; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974).

The HSCL-25 is a short-version of the original 90-item Hopkins Symptoms Checklist (HSCL) designed for quick assessment of general anxiety and depression. It is

comprised of 25-items: 10-items anxiety, 15-items depression. Each question is scaled along four response categories (0 = Not at all, 1 = A little, 2 = Quite a bit, 3 = Extremely). Two scores are calculated: the total score is the average of all 25 items, while the depression score is the average of the 15 depression items. Studies of several populations have shown that total score is strongly correlated to severe emotional distress, and the depression score is correlated with the concept of depression as outlined in the DSM-IV (American Psychiatric Association, 1994). To date this scale has been translated to over a dozen languages, with high Cronbach's alpha's ($\alpha = .87 - .97$; Coyne, Palmer, Shapiro, Thompson, & DeMichele, 2004; Ichikawa, Nakahara, & Wakai, 2005; Miyazaki, Dewaraja, & Kawamura, 2006; Mouanoutoua & Brown, 1995).

Measures of Neuro-developmental Stability (Intelligence)

The Raven's Advanced Progressive Matrices 18-item (APM-18; Sefcek, Miller, et al., 2007).

This 18-item scale is a medium-length form of the 36-item Ravens Advanced Progressive Matrices (APM; Raven, Court, & Raven, 1998). It was developed for use in time-constrained and classroom settings, where the use of the complete 40-minute APM would be unreasonable. By choosing 1 of 8 possible answers, participants are asked to complete a series of 18 visual patterns in which a piece of each pattern missing. Items are arranged in order of increasing difficulty with a time limit of 25-minutes. This scale has shown moderate internal consistency ($\alpha = .71 - .79$), and convergent validity with other measures of intelligence (e.g. the Mill-Hill Vocabulary Scale, The Shipley Institute of Living Scale, SAT, ACT, GPA, and measured of creativity).

The Mill-Hill Vocabulary Scale Senior Multiple Choice Form, Sets A & B (MHV-MC) (Raven, Raven, & Court, 1997).

The MHV-MC is a 2-set, 68-item, multiple-choice vocabulary test developed as a measure of reproductive (storage and retrieval) ability to be used in conjunction with the Ravens Matrices tests. The purpose of this scale is to record one's ability to recall acquired information. To this extent it is related to one's education, cultural background, and familiarity with the administration language. Published reliabilities of this scale are high with test-retest reliabilities between .87 - .95, and split-half reliabilities over .90.

The Shipley Institute of Living Scale (SILS) (Zachary, 1986).

The SILS is comprised of two timed (10-minute) subscales; vocabulary (40-items), and abstraction (20-items). This test is predominantly used in English-speaking countries and is appropriate for average individuals from 14 years to adult, who are motivated test-takers. Interpretation of the score can be either based on individual subscales or on a Total SILS score [Subscale 1 + (Subscale 2*2)]. Published norms are from 322 army recruits with subscale split-half reliabilities reported as .87 (Vocabulary), .89 (Abstraction), and .92 (Total) (Zachary, 1986).

Academic Performance.

Academic performance was measured through two self-report measures; grade-point average (GPA) and Scholastic Aptitude Test (SAT) scores. A variety of studies have identified moderate to strong correlations between these measures and a variety of other cognitive abilities including; intelligence, personality, physical health, and psychopathology (Barton, Dielman, & Cattell, 1971; Brown, 1994; Dyer, 1987; Mouw &

Khanna, 1993)

Statistical Analyses

All statistical analyses were conducted using SAS version 8.2 for Windows (SAS Institute, 1999). Cronbach's alphas and bivariate correlations were computed using the *PROC CORR* procedure. Exploratory Factor analyses were conducted by using the *PROC FACTOR* procedure. Parallel analyses were conducted through use of the *PROC IML* procedure. Multiple Regression analyses aimed at testing each of the theoretical models were conducted using *PROC REG*.

Results

Scale Reliabilities

Internal consistency estimates were calculated for each scale through the use of the Cronbach's alpha coefficient. In all but one scale the traditionally applied cut-off for reasonable internal consistency reliability ($\geq .70$) was achieved (Nunnally, 1978; see Table 2, for complete list of scale reliabilities). However, the alpha coefficient for the negative symptoms scale was low ($\alpha = .59$). This is not surprising as a high internal consistency of this scale would imply a high comorbidity of negative symptoms. As the mean age of this college sample is 19.11 years, it would be improbable, if not unfortunate, for individuals to show a multitude of these negative symptoms.

Due to researcher assistant turnover 10-different individuals made the facial asymmetry and body asymmetry measurements, making assessment of the interrater reliabilities impossible. Therefore, the results reported in relation to these measures of asymmetry may be treated as subject to measurement bias.

Common Factor Models

Previous research utilizing the *K-Factor* and *F-Factor* (denoted *Covitality*) used theoretically derived standardized regression coefficients to define factors (see Figueredo, Vasquez, Brumbach, et al., 2006, for review). These factors were determined by constructing each factor based on the correlation of each scale to a theoretically defined factor model. To determine how these data fit the theoretical models, exploratory factor analyses (EFA) were conducted for each of the hypothesized higher-order factors (*K-Factor*, *F-Factor*, and *g-factor*).

Exploratory K-Factor Loadings

First, a principal factors procedure with promax rotation was used as an exploratory measure of the *K-Factor*. Two eigenvalues (1.66 and 1.53) meeting the subjective scree test indicated that a two factor (rather than the theoretically predicted single-factor) solution best fit these data (Gorsuch, 1983). To determine if the two factor solution was indeed the best “objective” solution a parallel analysis procedure was conducted (PROC IML: SAS Institute, 1999). The parallel analysis (Horn, 1965) estimated the sampling distribution of the eigenvalues with a Monte Carlo procedure. A large number (e.g. 1000) of random data sets is created with the same sample size and the same number of items as the real data set. Observed factors with eigenvalues exceeding the statistically significant eigenvalues of the estimated sampling distribution were retained. This analysis indicated that two factors should be retained.

Shown in Table 3a are the factor loadings, for this two-factor solution. These factors accounted for 41% of the unique and 85% of the common variance, with a factor

intercorrelation of -.06. Factor 1 (accounting for 44% of the common variance) is comprised of all the “prosocial” scales theoretically associated with the *K-Factor* (i.e. Mini-*K*, parental investment, family and friend support, altruism, and romantic attachment), and has thus been named *K-Factor*. Factor 2 (accounting for 41% of the common variance) is comprised of scales that are associated with self-endorsed attitudes related to psychopathy (i.e. psychopathic personality, mating effort, and risk taking), and has therefore been named *psychopathic attitudes*. A Varimax rotation retained these factors with the same scales loading on each. The similarities between the orthogonal and oblique rotation suggested that the orthogonal solution offered a good fit for these data (Pedhazur & Schmelkin, 1991).

Due to the results of these analyses, the psychopathic personality inventory, MES, and risk-taking measure were dropped as indicators of *K* from subsequent analysis utilizing the *K-Factor*. Rather, they were included in subsequent analyses as a *Psychopathic Attitudes* factor (*P-Factor*).

Exploratory F-Factor Loadings

An exploratory principle factors procedure with promax rotation was used to explore the *F-Factor*. One eigenvalue (3.69) meeting the subjective scree test indicated that a one factor solution best fit these data. A parallel analysis was conducted to objectively test this decision. This parallel analysis also indicated that one factor should be retained.

Shown in Table 3b are the factor loadings, for this single-factor solution. This factor accounted for 36% of the unique and 76% of the common variance. Due to the low

factor loadings of three items (BMI, FFA, and BFA) they were dropped from subsequent data analysis. Retained in this factor are seven items which loaded on the factor in their theoretically predicted way positive health aspects: general health, SF-36, and SWB, and negative aspects: anxiety, depression, BDI, and negative health symptoms.

Exploratory g-factor Loadings

The general intelligence, *g-factor* was explored through a principal factors procedure. One eigenvalue (1.70) was retained through the use of the subjective scree test. A parallel analysis likewise indicated that one factor should be retained.

Shown in Table 3c are the factor loadings for this *g-factor*. This factor accounted for 23% of the unique and 70% of the common variance. Overall this factor represents a strong *g-factor* with positive factor loadings on all measures of intelligence.

Unit-weighted Higher Order factors

Due to the results of the various EFA's, four higher order factors were created by defining the factors based on direction of their factor loadings from the EFA's. These were the; *K-Factor*, *P-Factor*, *F-Factor*, and *g-factor*. Shown in Table 4 are the factor loadings (reported as standardized regression coefficients) for each of the items comprising the latent constructs. These factors were the ones used in all subsequent analyses.

Bivariate correlations among the higher-order factors of interest were calculated by the PROC CORR procedure (SAS Institute, 1999) and are displayed in Table 5a. As can be seen, the only significant relationship to emerge was between the *K-Factor* and *F-Factor* ($r = .31, p < .001$). However, as with the pilot data, this relationship is weaker

than with previous samples (e.g. Figueredo et al., 2004). As an exploration of this finding, the correlation between each pair of factors was calculated separately for males and females. Shown in Table 5b are these correlations. Again, the only significant correlation to emerge was between the *K-Factor* and *F-Factor*, but only for females ($r_{female} = .45, p < .001$; $r_{male} = .14, p > .05$). This suggests that at least for males, there is no relationship between life-history and health. Theoretical implications of this will be explored below.

In order to further explore the lack of correlation between either the *K-Factor* or *F-Factor* and the *g-factor* bivariate correlations were calculated in relation to each of the specific measures comprising the *g-factor*. Shown in Table 6 are these correlations. As can be seen, GPA was significantly related to both the *K-Factor* and *F-Factor* ($r = .15, p < .05$; and $r = .17, p < .05$, respectively). This correlation, however is not impressive and no other significant correlations were identified in relation to either the *K-Factor* or *F-Factor*.

Multiple Regression Models

Multiple Regression Analyses (PROC REG) were conducted to explore each of the alternative theoretical models. Because there is a large body of literature supporting sex differences in each of the higher-order factors of interest; life-history strategy, psychopathic attitudes and behavior, general health, and intelligence, the sex variable was entered as a predictor in each model. Each test was conducted so that the criterion variable was predicted by a higher order factor, sex, and the higher order factor by sex interaction. Interaction terms were named *K-sex*, *P-sex*, *F-sex*, and *g-sex*. Results,

reported as one least-squares (OLS) β -weights, their associated t tests, and significance levels, are reported in Table 7.

Tests of the Differential-K Model

The first regression analysis examined the relationship between the *K-Factor* and *P-Factor*, as predicted by the *Differential-K* life history model. The *P-Factor* was used as the criterion variable with the ordered predictors of *K-Factor*, sex, and the *K-sex* interaction. The overall model was significant ($R^2 = .12$; $F = 8.83$, $p < .001$), with the only significant predictor of this model being sex (t test = 5.14, $p < .001$). This is further support for the results of the EFA leading to a two-factor solution for the theoretically derived *K-Factor*. The fact that sex predicts the *P-Factor* is not surprising as there is a large body of literature showing that males exhibit both more psychopathic behavior and attitudes (see Hare, 1993, for review).

The second regression analysis examined the prediction that life-history would drive somatic effort, as measured by *K-Factor* $\rightarrow$ *F-Factor* relationship. The *F-Factor* was entered as the criterion variable with the ordered predictors of *K-Factor*, sex, and the *K-sex* interaction. The overall model was significant ($R^2 = .13$; $F = 9.27$, $p < .001$), as were the *K-Factor* and *K-Factor* x sex interaction terms (t test = 4.90, $p < .001$; and $F = 2.18$, $p < .05$, respectively). This model suggests that while life history is itself a significant predictor of somatic effort, there is something special about females in this relationship. Theoretical reasons for this finding are explained below in relation to Wilson's (1957) theory of antagonistic pleiotropy.

The third regression analysis tested the prediction that life-history strategy is predictive of general intelligence. The *g-factor* was entered as the criterion variable, with *K-Factor*, sex, and the *K-sex* interaction. This model ($R^2 = .00$; $F \leq 0.77$, $p > .05$) and all individual predictors were non-significant (see Table 7). This is a further confirmation that the *Differential-K* model does not seem to account for general intelligence, as measured by the *g-factor*.

Tests of the General Fitness Model

Two regression analyses were conducted to explore Miller's (2000) general fitness model. The first one examined the proposition that the *F-Factor* would be predictive of life-history. Here the *K-Factor* was entered as the criterion variable, with the ordered predictor variables being *F-Factor*, sex, and the *F-Factor* x sex interaction. The overall model was significant ($R^2 = .13$; $F = 9.29$, $p < .001$), as were the *F-Factor* and *F-Factor* x sex interaction terms (t test = 5.02, $p < .001$; and $F = -2.40$, $p < .05$, respectively). This analysis suggests that while general fitness is related to higher *K-Factor* scores, there is a stronger effect for females than males.

The second regression analysis for the general fitness model examined the prediction that phenotypic fitness would predict general intelligence. The *g-factor* was entered as a criterion variable in order to examine the relationship between the *F-Factor*, sex, and *F-sex* interaction. This model ($R^2 = -.02$; $F = 0.59$, $p > .05$) and all individual predictors were found to be non-significant (see Table 7). This is a further confirmation that the general fitness model does not account for variation in general intelligence.

Test of the Hybrid Model

A regression analysis was conducted to examine if life-history strategy and general fitness jointly account for the variance in human general intelligence. The *g-factor* was defined as the criterion, with *K-Factor*, *F-Factor*, *K-Factor* x *F-Factor*, *sex*, *K-sex*, *F-sex*, and the *K-Factor* x *F-Factor* x *sex* interactions as the predictors. This model was not significant ($R^2 = .02$; $F = .63$, $p > .05$), nor were any of the predictor variables (See Table 7). These results suggest that life-history strategy and general fitness do not jointly account for general intelligence.

Discussion

Using the method of multiple working hypotheses this study was developed to examine several predictions stemming from Rushton's (1999) *Differential-K theory*, Miller's (1999) general fitness model, and a complementary hybrid model. Primarily, these predictions examined the relationships among the latent constructs of the *K-Factor*, *F-Factor*, and *g-factor*, with secondary analyses examining the subcomponents comprising each factor. Results here suggest only minimal support for any of the primary predictions. Specifically, there was a significant relationship between the *K-Factor* and *F-Factor*. However, even this relationship was less robust than previous findings (see Figueredo, et al., 2004). Other results, indicated no relationship between either the *K-Factor* and *g-factor*, the *F-Factor* and *g-factor*, or *K-Factor-F-Factor* interaction and *g-factor*. Explanations for these findings are presented below. Where applicable, (based on similar theoretical explanations) predictions have been grouped together.

Outcomes of Specific Predictions

Prediction 1

Prediction 1 examined the contention that behaviors deemed “socially deviant” or delinquent in nature, should have a negative relationship to the *K-Factor* (load negatively on *K*). In this study, three scales (Mating Effort Scale, Risk Taking Questionnaire, and The Psychopathic Personality Inventory) were used to measure this prediction. Rather than showing a single robust *K-Factor* with negative loadings on these scales, an EFA revealed two uncorrelated factors ($r = -.06, p > .05$) accounting for 85% of the common variance. The first factor retained the prosocial items of the original *K-Factor*, and has thus retained its name. Due to the nature of the questions contained in the three other scales, the second factor has been named “psychopathic attitudes” (or *P-Factor*).

Research on life history has revealed that a long-list of behaviors deemed socially unacceptable and deviant are positively related to a low-*K*, fast reproductive strategy (as well as being negatively related to a high-*K*, slow strategy). For example, Charles and Egan (2004) found that mating effort (as measured by the MES) was positively related to self-reported delinquency in a large sample of adolescents; while Egan, et al. (2005) found a relationship between Mating effort and sensational, or macabre interests. Other researchers have identified a relationship between delinquency and promiscuity (Rowe & Rodgers, 1989), and mating effort (Rowe, Vazsonyi, & Figueredo, 1997). These findings show correlations among different single measures of life history (e.g., mating effort) and what has been considered as socially deviant behavior.

However, consistent with the current results, Brumbach (2007) found that social deviance, as a latent factor, was not related to life-history in a large sample of nationally represented young adults. She suggested that young adults may not have not yet

incorporated socially deviant behaviors into their life history strategy, but that they might do so by middle adulthood. This argument may be based on the concept of *heterochrony*, which refers to the differential timing in the maturation of phenotypic traits that lead to changes in sizes and shape. For example, an abundance of testosterone in young males influences them to engage in high levels of mating effort, which may lead to risky displays (Wilson & Daly, 1985). At the same time the brain's frontal lobes have not yet fully myelinated, which results in lowered inhibition of such behaviors. As males age frontal lobe white matter increases and may continue into their fifth decade (Bartzokis, Beckson, Lu, Nuechterlein, Edwards, & Mintz, 2001). Such delayed development may be a partial cause of the risk-prone behavior of younger versus older men, and a general “mellowing” of behavior as they age..

Two alternative explanations might account for such findings; i.e. the construct validity of the measures used, and the sample demographics. Brumbach's (2007) study, may suffer from the former as self-report questionnaires aimed at assessing both social deviance and delinquency were employed. These questions asked for participants to endorse how many times they had engaged in a behavior in the most recent 12-month period. It is possible that such a short timeframe gave a narrowed view of their behaviors. It is also possible that participants over or under reported their behaviors. To our knowledge no lie scale or social desirability scale was incorporated into the study.

The current study however, may suffer from both the problem of construct validity and sample demographics. The latter is most relevant here due both to the young age of a university-level sample. It is likely that young adults who are attending a state

university are a self-selected group and are not an accurate representation of evolved tendencies across the entire human species (Sefcek, et al., 2006).

Further, the scales used here were not measures of behavior. Rather they were measures of attitudes, asking questions about how one believes they act, or would act, in a specific context. A long history of research spanning back several decades has shown that attitudes do not necessarily predict behavior (e.g. LaPiere, 1934; Wicker, 1969), unless they are strongly held (Bassili, 1995) and vitally affect the holder's interest (Sivacek & Crano, 1982). In this sample, it is unlikely that any of the participants are extremely delinquent, psychopathic, or socially deviant. Therefore, simply because they may have attitudes directed towards social deviance, their behavior may not follow suit. Further research needs to be conducted using a more representative sample of humans.

Predictions 2 & 4

Because each of these predictions were concerned with the relationship between the *K-Factor* (i.e. life history strategy) and *F-Factor* (*Covitality*; somatic effort) they will be examined together. As with previous studies (e.g. Figueredo, et al., 2004), a significant positive relationship between these factors was found. This relationship is however, weaker than previously reported. Further, when examined as a function of sex, the relationship between these factors disappears for males, but becomes stronger in females (on par with previous findings). Two complementary explanations may account for these findings. First, due to the young age of this sample there may simply be very little phenotypic variation in general health and psychopathology. Such variation may become more apparent with increasing age. This coupled with the sex differences noted here are

not surprising when we look at the *theory of antagonistic pleiotropy* (Williams, 1957).

The *theory of antagonistic pleiotropy* (Williams, 1957) refers to the fact that many genes producing fitness-enhancing effects while an organism is young have detrimental effects on survival and longevity. The fitness-enhancing effects of any such character are therefore age-dependent and will decrease the longer the character is expressed (Williams & Day, 2003). This tends to lead to selection operating more strongly on individuals when they are at or near peak reproductive age (younger) at a cost to overall longevity. Hence younger individuals will tend to be healthier than older individuals.

If low-*K* individuals invest more resources in reproductive effort at an earlier age, there may be a difference in age of peak reproductive effort in comparisons of low and high-*K* individuals. As such, we might expect to find very healthy, robust, low-*K* individuals when young, but see a rapid decline in this robustness as they mature. Across their lifespan there should be higher rates of intrinsic disease and mortality, and we should note an earlier onset and more rapid degradation of health when compared to those who are higher-*K*. This mortality may come from a compromised immune system, a breakdown of some somatic maintenance mechanism (e.g., cellular turnover), or the negative effects of sex hormones associated with antagonistic pleiotropy (e.g., testosterone). Despite this prediction however, the current sample of 18-year olds might be too young to identify such variation.

Antagonistic pleiotropy also predicts that the sex which engages in higher mating effort will have encountered stronger selection when young, compared to the sex which engages in lower mating effort. In humans, males are the higher mating-effort sex. As

such, human males should show lower deleterious effects when young, at a cost for later health and longevity. One example of such an effect can be illustrated with testosterone.

Testosterone is a male sex hormone that increases musculature and aggressiveness. This may have a fitness-enhancing effect when young as it increases competition for mates. However, it also shows many negative effects from prolonged, or heightened, exposure. These negative effects include immune suppression and increased susceptibility to prostate cancer (leading to poor health), as well as aggressiveness, recklessness, and risk-taking that may lead to injury or death (Wilson & Daly, 1985), and may even be related to the 7-year lower life expectancy of human males compared to females (Williams & Nesse, 1991). Simply put, selection operates on the young, at a cost to the old; and selection operates more harshly on human males than females (as the current rate of male-pattern baldness surely shows).

More research needs to examine this proposition in relation to individual differences in human life-history strategy. Specifically, research should examine if the relationship between the *K-Factor* and *F-Factor* reappears for older males. Further, research should examine if there is a cross-over age, where the relationship between the *K-Factor* and *F-Factor* becomes more pronounced in males than in females.

Unfortunately, the regression analysis is unable to determine the causal direction of the *K-Factor/F-Factor* relationship.

Predictions 3, 5, & 6

Predictions 3, 5, and 6 examined the relationship between general intelligence, as measured by the *g-factor*, and the *K* and *F-Factors*. Counter to the predictions there were

no direct relationships between either K or F and 'g', or a $K \times F$ interaction and 'g'. This is, at first, surprising due to the large amount of literature in the realm of intelligence showing a relationship to many key life history (e.g., birth weight, Richards, Hardy, Kuh, & Wadsworth, 2001; and longevity, Whalley & Deary, 2001) and health traits (Gottfredson & Deary, 2001). Further, other research (e.g., Smith, 1989) has established the connection between brain size and intelligence and shown that brain weight is related to a variety of life history traits (e.g. life span, length of gestation, age of weaning, age at sexual maturity, interbirth interval, and body weight). Although it is intuitive to extrapolate from such research that intelligence and life-history strategy will be linked, alternative hypotheses should also be taken into account

First, while it has been established that there in humans there is a moderate relationship between both head size and intelligence (mean $r = .19$) and brain size based on MRI and intelligence ($r = .40, p < .05$; Wickett, Vernon, & Lee, 1994), much of the variance in intelligence is still unaccounted for by either metric. Arguments using head/brain size/weight as a proxy for intelligence may introduce a high-degree of measurement error. Further, whereas most research reports single item correlations between variables (e.g. brain size and score on a single intelligence measure), the current study was concerned with correlations between higher-order factors comprised of a variety of measurements. If, as is argued, both life history strategy and general intelligence are composite entities, inclusive of a variety of genetically and phenotypically linked traits, then results based on higher-order factor modeling, such as

in the current study, might be more realistic than results examining bivariate correlations between two separate traits.

A second explanation for such findings stems from the difficulty in extrapolating within species comparisons based on between species differences. Rushton (2004) identified positive relationships between brain weight and many life history traits (e.g., gestation time, birth weight, lactation time, body size, age of first mating, and longevity) when comparing between species. However, other research has shown that when comparing within species differences, factors such as body size is related to large litters, early maturation, high reproductive rates, and low adult mortality (Altmann & Alberts, 2003). Disparate findings such as these suggest that making claims of within species variation based on between species comparisons might be in error.

Additionally, the theoretical nature of intelligence needs to be addressed. For example, *Differential-K theory* posits that general intelligence was shaped by selection to solve problems associated with newly encountered harsh, but predictable, ecological conditions. It would seem that in such an ecological context selection would be better served to create inflexible, instinctual mechanisms based on such predictability. Other theories of general intelligence tend to focus on novel adaptive problems as being the drive for its evolution (e.g., Kanazawa, 2003). If true, this might explain why there is no difference in the *g-factor* between high-and-low-*K* individuals, as novel encounters are likely to occur across ecological niches.

Perhaps that while general intelligence may not be a product of life-history strategy, social intelligence is. Social intelligence deals with behaviors such as social

technique, “knowledge of social matters, susceptibility to stimuli from other members of a group, and insight into the temporary moods and underlying personality traits of strangers”, which together aid an individual in getting along with others, in general (Vernon, 1933, p.44). The larger social group size, higher group cohesion and altruism, yet more intersexual aggression thought to be associated with *K*-selection might provide a selective pressure for the ability to be able to gauge and react correctly to social situations, i.e. socially intelligent.

Other cognitive capacities may also be related to life-history strategy. As Pianka (1970) argued, the *r/K* dichotomy of life history is an adaptive solutions to problems associated with the predictability and stability of environments. Many cognitive capacities associated with executive function may therefore be better candidates as indicators of life-history. For example, the abilities to update information, inhibit behavior, and shift between multiple tasks are each components of executive function. Because a harsh, predictable, *K*-selected environment, would necessitate the ability to delay gratification, plan ahead, and live in large social groups; high-*K* individuals should have the ability to control impulsive behaviors, shift attention in a quick, flexible manner, and integrate, organize, and monitor their performance. These abilities have been shown to be unrelated to general intelligence (Friedman, Miyake, Corley, Young, DeFries, & Hewitt, 2006). Indeed, in a recent unpublished study, MacDonald, Figueredo, & Wenner (2007) have reported a significant positive correlation between the *K-Factor* and an executive function factor comprised of such measures, with neither one being correlated to general intelligence.

Many previous studies report a relationship between measures of general intelligence and single proposed indicators of fitness such as; body symmetry (Furlow, Armijo-Prewitt, Gangestad, & Thornhill, 1997; Prokosch, Yeo, & Miller, 2005), physical attractiveness (Zebrowitz, Hall, Murphy, & Rhodes, 2002), general health (Jensen & Sinha, 1993), and longevity (Whalley & Deary, 2001). The current data are however, inconsistent with these findings.

This is a difficult issue to resolve as the available literature examining the proposed “fitness factor” has been met with positive results, i.e. there is a strong relationship between single measures of morphodevelopmental stability (i.e. body symmetry, hormone markers) and neurodevelopmental stability (i.e. general intelligence, psychopathology, personality). The only explanation offered here is that again, while previous studies were utilizing single-trait indicators of these constructs, the current study used higher-order factors comprised of many different indicators. It may be that not all proposed indicators are equally representative of the proposed construct, and therefore inclusion in the construct mediates this relationship. Further work is necessary to resolve this issue.

Limitations and Future Directions

There are several limitations associated with the current study. First, as noted above, the age of this sample may not be a representative age for exploring the evolutionary psychology of human mating behavior (Sefcek, et al., 2006). Because both fertility and mating effort peak in the mid-to-late 20's, looking for variation in mating behavior among 18-year old college students might be in error. Additionally, the college

environment is likely to be very different than what was experienced over the course of both the developmental history of the student, as well as the evolutionary history of humans. College students experience novel stimuli (e.g., little direct familial contact and supervision, large groups of unknown same-sex and other sex peers) which may necessitate both a restructuring of already learned strategies, but also an adoption of new ones. Future research needs to incorporate samples that are more varied in ages, cultures, and developmental histories.

A second limitation concerns the structure of the *K-Factor* itself. *Differential-K theory* predicts the use of a coherent behavioral strategy in the execution of one's life-history. Associated with this strategy should be a variety of attitudes, behaviors, and physiological features. To date, the measures used in the Arizona Life History Battery to define the *K-Factor* have focused on self-reports of recalled past behaviors, attitudes concerning their past behavior, and attitudes and predictions of how participants claim they will behave in future situations. For this *K-Factor* to be a valid indicator of life-history it will need to first, incorporate more objective measures of past, current, and future behavior; and second, incorporate more direct measures of reproductive behaviors (e.g., counting babies, counting past/current number of sexual partners). Research will also need to include observational data collection focusing on different aspects of theoretically important behaviors, e.g., do those who are measured as being low-*K* actively seek out more sexual partners than those who are measured as high-*K*?

REFERENCES

- Altmann, J., & Alberts, S. C. (2003). Intraspecific variability in fertility and offspring survival in a non-human primate: behavioral control of ecological and social sources. In K. W. Wachter & R. A. Bulatao (Eds.), *Biodemography of Fertility and Family Behavior*. (pp. 140-169). National Academy Press: Washington, DC.
- Ambrosini, P. J., Metz, C., Bianchi, M. D., Rabinovich, H., & Undie, A. (1991). Concurrent validity and psychometric properties of the Beck Depression Inventory in outpatient adolescents. *Journal of the American Academy of Child Adolescent Psychiatry*, 30, 51-57.
- American Psychiatric Association (1994). (DSM-IV) *Diagnostic and statistical manual of mental disorders*, (4th edition). Washington, DC: American Psychiatric Press, Inc.
- Arthur, W. Jr., & Day, D. V. (1994). Development of a short form for the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement*, 54, 394-403.
- Arthur, W., Jr., Tubre, T. C., Paul, D. S., & Sanchez-Ku, M. L. (1999). College-sample psychometric and normative data on a short form of the Raven Advanced Progressive Matrices Test. *Journal of Psychoeducational Assessment*, 17, 354-361.
- Barton, K., Dielman, T. E. & Cattell, R. B. (1971). The prediction of school grades from personality and IQ measures. *Personality*, 2(4), 325-333.
- Bartzokis, G., Beckson, M., Lu, P. H., Neuchterlein, K. H., Edwards, N., & Mintz, J. (2001). Age-related changes in frontal and temporal lobe volumes in men: a magnetic resonance imaging study. *Archives of General Psychiatry*, 58, 461-465.
- Bassili, J. N. (1995). Response latency and the accessibility of voting intentions: What contributes to accessibility and how it affects vote choice. *Personality and Social Psychology Bulletin*, 21, 686-695.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory* (2nd ed. Manual). San Antonio, TX: The Psychological Corporation.
- Belsky, J., Steinberg, L., & Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child Development*, 62, 647-670.
- Bock, G., Goode, J., & Webb, K. (2000), *The nature of intelligence*. Novartis Foundation Symposium 233. New York: John Wiley & Sons.

- Brim, O. G., Baltes, P. B., Bumpass, L. L., Cleary, P. D., Featherman, D. L., Hazzard, W. R., et al. (2000). National survey of midlife development in the United States (MIDUS), 1995–1996 [Computer file]. ICPSR version. Ann Arbor, MI: DataStat, Inc./Boston, MA: Harvard Medical School, Department of Health Care Policy [producers], 1996. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].
- Brown, M. W. (1994). Cognitive, interest, and personality variables predicting first-semester GPA. *Psychological Reports*, 74(2), 605-606.
- Brumbach, B. H. (2007). *Effects of harsh and unpredictable environments in adolescence on development of life history strategies: A longitudinal test of an evolutionary model*. Unpublished doctoral dissertation, University of Arizona, USA.
- Brumbach, B. H., Figueredo, A. J., & MacDonald, K. (2005). A constructive replication of the Super-K factor using the “Mini-K” short form. Paper. In Figueredo, A. J., (Chair), *The Psychometrics and Behavioral Genetics of Life History Strategy*. Annual Meeting of the Human Behavior and Evolution Society, Austin, Texas.
- Calle, E. E., Thun, M. J., Petrelli, J. M., Rodriguez, C., Heath, C. W. (1999). Body-mass index and mortality in a prospective cohort of U.S. adults. *New England Journal of Medicine*, 341(15), 1097-105.
- Chamberlin, T. C. (1890, 1965). The method of multiple working hypotheses: *Science*, 148, 754-759.
- Chapman, A. L., Gremore, T. M., and Farmer, R. F. (2003). Psychometric analysis of the Psychopathic Personality Inventory (PPI) with female inmates. *Journal of Personality Assessment*, 80, 164-172.
- Charles, K. E., & Egan, V. (2005). Mating effort strongly predicts self-reported delinquency in a normal adolescent sample. *Personality and Individual Differences*, 38, 1035–1045.
- Chisholm, J. S. (1996). The evolutionary ecology of attachment organization. *Human Nature*, 7, 1–38.
- Coyne, J. C., Palmer, S. C., Shapiro, P. J., Thompson, R., DeMichele, A. (2004). Distress, psychiatric morbidity, and prescriptions for psychotropic medication in a breast cancer waiting room sample. *General Hospital Psychiatry*, 26(2), 121-128.
- Craven, J. L., Rodin, G. M., & Littlefield, C. (1988). The Beck Depression Inventory as a screening device for major depression in renal dialysis patients. *International Journal*

- of Psychiatry and Medicine*, 18, 365-374.
- Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self report symptom inventory. *Behavioral Science*, 19, 1 – 15.
- Dyer, E. D. (1987). Can university success and first-year job performance be predicted from academic achievement, vocational interest, personality and biographical measures? *Psychological Reports*, 61(2), 655-671.
- Edens, J. F., Poythress, N. G., & Lilienfeld, S. O. (1998, August). *Validation of the Psychopathic Personality Inventory in correctional and community samples*. Paper presented at the 106th annual conference of the American Psychological Association, San Francisco.
- Egan, V., Figueredo, A. J., Wolf, P., McBride, K., Sefcek, J., Vasquez, G., and Charles, K. (2005) Sensational interests, mating effort, and personality: Evidence for cross-cultural validity. *Journal of Individual Differences*, 26, 11-19.
- Ellis, L. (1988). Criminal behavior and r/K selection: an extension of gene-based evolutionary theory. *Personality and Individual Differences*, 9, 697–708.
- Ellis, B. J. (2004). Timing of pubertal maturation in girls: An integrated life history approach. *Psychological Bulletin*, 130, 920–958.
- Figueredo, A. J., Sefcek, J., Vasquez, G., Brumbach, B. H., King, J. E. & Jacobs, W. J., (2005). Evolutionary theories of personality. In Buss, D.M., (Ed.), *Handbook of Evolutionary Psychology* (pp. 851-877). Hoboken, NJ: John Wiley & Sons.
- Figueredo, A. J., Vásquez, G., Brumbach, B. H., & Schneider, S. M. R., 2007. The K-Factor, covitality, and personality: A psychometric test of life history theory. *Human Nature*, 18(1), 47-73.
- Figueredo, A. J., Vásquez, G., Brumbach, B. H., & Schneider, S. M. R. (2004). The heritability of life history strategy: The K-Factor, covitality, and personality. *Social Biology*, 51, 121-143.
- Figueredo, A. J., Vásquez, G., Brumbach, B. H., Schneider, S. M. R., Sefcek, J. A., Tal, I. R., Hill, D., Wenner, C. J., & Jacobs, W. J. (2006). Consilience and life history theory: From genes to brain to reproductive strategy. *Developmental Review*, 26, 243-275.
- Figueredo, A. J., Vásquez, G., Brumbach, B. H., Sefcek, J. A., Kirsner, B. R., & Jacobs, W. J. (2005). The *K-Factor*: Individual differences in life history strategy.

- Personality and Individual Differences*, 39, 1349-1360.
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, 17, 172-179.
- Gadgil, M., & Solbrig, O.T. (1972). The concept of r and K selection: Evidence from wild flowers and some theoretical considerations. *American Naturalist*, 106, 14-31.
- Galton, F. (1869, 1962). *Hereditary Genius: An Inquiry into its Laws and Consequences*. London: Macmillan/Fontana.
- Gangestad, S. W., & Thornhill, R. (2003). Facial masculinity and fluctuating asymmetry. *Evolution and Human Behavior*, 24, 231-241.
- Garrow, J. S., & Webster, J. (1985). Quetelet's index (W/H^2) as a measure of fatness. *International Journal of Obesity*, 9, 147-153.
- Gorsuch, R. L. (1983). *Factor Analysis*. Hillsdale, N.J.: Erlbaum.
- Gottfredson, L. S., & Deary, I. J. (2004). Intelligence Predicts Health and Longevity, but Why? *Current Directions in Psychological Science*, 13(1), 1-4.
- Hare, R. D. (1993). *Without conscience: The disturbing world of the psychopaths among us*. New York: Guilford Press.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185.
- Houle, D. (1991). Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. *Evolution*, 45, 630 – 648.
- Ichikawa, M., Nakahara, S., & Wakai, S. (2005). Lowered tuberculosis notifications and deterred health care seeking during the sars epidemic in Hong Kong. *American Journal of Public Health*, 95, 933-4.
- Jacobson, A. Steklis, H.D., Steklis, N.G., LeClair, & Fawcett, J.K.(2004). Exploring the Application of Human Faical Fluctuating Asymmetry Methodology in a Great Ape. Paper. In Steklis, H.D. *New Approaches to Facial Attractiveness* (Chair). Symposium. Annual Meeting of the Human Behavior and Evolution Society, Berlin, Germany.
- Jensen, A. (2000). The g Factor: Psychometrics and Biology. In G. Bock, J. Goode, & K.

- Webb (Eds.), *The nature of intelligence*. Novartis Foundation Symposium 233. New York: John Wiley & Sons, pp. 37-47.
- Jensen, A. R., & Sinha, S. N. (1993). Physical correlates of human intelligence. In P. A. Vernon (Ed.), *Biological approaches to the study of human intelligence*. Norwood, NJ: Ablex.
- Kanazawa, S. (2003). General intelligence as a domain-specific adaptation. *Psychological Review*, 111, 1761 – 1776.
- Kirk, K. M., Blomberg, S. P., Duffy, D. L., Heath, A. C., Owens, I. P. F., & Martin, N. G. (2001). Natural selection and quantitative genetics of life-history traits in western women: A twin study. *Evolution*, 55, 423–435.
- Krebs, C. J., Gaines, M. S., Keller, B. L., Myers, J. H. & Tamarin, R. H. (1973). Population cycles in small rodents. *Science*, 179, 35-41.
- LaPierre, R. T. (1934). Attitudes vs. Actions. *Social Forces*, 13, 230-237.
- Leggett, W.C., & Carscadden, J.E. (1978). Latitudinal variation in reproductive characteristics of American shad (*Alosa sapidissima*): evidence for population specific life history stages in fish. *Journal of the Fisheries Research Board of Canada*, 35, 1469-1478
- Levin, B. E., Llabre, M. M., Weiner, W. J. (1988). Parkinson's disease and depression: psychometric properties of the Beck Depression Inventory. *Journal of Neurology and Neurosurgical Psychiatry*, 51, 1401-1404.
- Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment*, 66, 488-524.
- Lovibond, P. F., Lovibond, S. H. (1995). The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33, 335-343.
- MacDonald, K.M., Figueredo, A.J., & Wenner, C. (2007). Life history strategy, executive functions, and personality. Paper. In Figueredo, A.J., (Chair), *Correlates of life history strategy*. Annual Meeting of the Human Behavior and Evolution Society, Williamsburg, Virginia.
- McArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*, Princeton, NJ: Princeton University Press.

- McHorney C. A., Ware, J. E., Lu, J. F., Sherbourne, C. D. (1994). The MOS 36-Item Short-Form Health Survey (SF-36): III. Tests of Data Quality, Scaling Assumptions, and Reliability Across Diverse Patient Groups. *Medical Care*, 32(1), 40–66
- Mealey, L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. *Behavioral & Brain Sciences*, 8(3), 523-599.
- Miller, G. F. (2000a). Mental traits as fitness indicators: Expanding evolutionary psychology's adaptationism. In D. LeCroy & P. Moller (Eds.), *Evolutionary approaches to human reproductive behavior. Annals of the New York Academy of Science*, 907, 62-74.
- Miller, G. F. (2000b). *The Mating Mind: How sexual choice shaped the evolution of human nature*. New York: Doubleday.
- Miller, G. F. (1998). How mate choice shaped human nature: A review of sexual selection and human evolution. In C. Crawford & D. Krebs (Eds.), *Handbook of evolutionary psychology: Ideas, issues, and applications* (pp. 87-129). Lawrence Erlbaum.
- Miyazaki, T., Dewaraja, R., & Kawamura, N. (2006). Reliability and validity of the scales related to post traumatic stress disorder of Sri Lanka version. *International Congress Series*, 1287, 82-85
- Mouanoutoua, V. L., & Brown, L. G. (1995). Hopkins Symptom Checklist-25, Hmong Version: A screening instrument for psychological distress. *Journal of Personality Assessment*, 64, 376-383.
- Mouw, J. T. & Khanna, R. K. (1993). Prediction of academic success: A review of the literature and some recommendations. *College Student Journal*, 27(3), 328-336.
- Murray, C. (1998). *Income Inequality and IQ*. Washington: AEI Press.
- National Institute of Health (2005). *Aim for a Health Weight*. Retrieved December 12, 2005 from http://www.nhlbi.nih.gov/health/public/heart/obesity/lose_wt/risk.htm
- Naughton, M. J., & Wiklund, I. (1993). A critical review of dimension-specific measures of health-related quality of life in cross-cultural research. *Quality of Life Resources*, 2, 397-432.
- Nunnally, J. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Oppenheimer, S. (2003). *The Real Eve: Modern Man's Journey Out of Africa*. Carrol & Graff Publishers, New York, New York.

- Pedhazur, E.J. & Schmelkin, L.P. (1991). *Measurement design, and analysis: An integrated approach*. Hillsdale, NJ. Erlbaum.
- Pianka, E. R. (1970). On *r*-and *K*-selection. *American Naturalist*, 104, 592-596.
- Pigliucci, M., & Hayden, K. (2001). Phenotypic plasticity is the major determinant of changes in phenotypic integration in *Arabidopsis*. *New Phytologist*, 152(3), 419-430.
- Platt, J. R. (1964). Strong inference. *Science*, 146, 347-353.
- Plomin, R. (2003). *Behavioral Genetics in the Postgenomic Era*. Washington, DC: APA Books.
- Poythress, N. G., Edens, J. F., & Lilienfeld, S. O. (1998). Criterion-related validity of the Psychopathic Personality Inventory in a prison sample. *Psychological Assessment*, 10, 426-430.
- Prokosch, M. D., Yeo, R. A., & Miller, G. F. (2005). Intelligence tests with higher *g*-loadings show higher correlations with body symmetry: Evidence for a general fitness factor mediated by developmental stability. *Intelligence*, 33, 203-213.
- Quetelet, Adolphe (1842). *A Treatise on Man and the Development of His Faculties*. Oxford University. Retrieved December 10, 2005, from http://books.google.com/books/pdf/A_treatise_on_man_and_the_development_of.pdf
- Rodgers, J. L., Hughes, K., Kohler, H.-P., Christensen, K., Doughty, D., Rowe, D. C., et al. (2001). Genetic influence helps explain variation in human fertility outcomes: Evidence from recent behavioral and molecular genetic studies. *Current Directions in Psychological Science*, 10, 184-188.
- Raven, J. C., Court, J. H., & Raven, J. (1998). *Manual for Raven's Standard Progressive Matrices (1998 edition)*. Oxford, England: Oxford Psychologists Press.
- Raven, J, Raven, J. C. & Court, J. H. (1997). *Mill Hill Vocabulary Scale: 1998 Edition*. Oxford: Oxford Psychologists Press.
- Richards, M., Hardy, R., Kuh, D., & Wadsworth, M. E. J. (2001). Birth weight and cognitive function in the British 1946 birth cohort: longitudinal population based study. *British Medical Journal*, 322, 199-203.
- Rowe, D. C. (2000). Environmental and genetic influences on pubertal development:

- Evolutionary life history traits? In J. L. Rodgers, D. C. Rowe, & W. B. Miller (Eds.), *Genetic influences on human fertility and sexuality: Recent empirical and theoretical Findings* (pp. 147–168). Boston: Kluwer.
- Rowe, D. C., & Flannery, D. J. (1994). An examination of environmental and trait influences on adolescent delinquency. *Journal of Research in Crime and Delinquency*, 31, 374–379.
- Rowe, D. C., & Rodgers, J. L. (1989). Behavior genetics, adolescent deviance, and “d”: Contributions and issues. In G. Adams, R. Montemayor, & T. Gullotta (Eds.), *Advances in Adolescent Development* (pp. 38–70). Newbury Park, CA: Sage.
- Rowe, D. C., Vazsonyi, A. T., & Figueredo, A. J. (1997). Mating effort in adolescence: Conditional or alternative strategy? *Personality and Individual Differences*, 23(1), 105–115.
- Rowe, L., & Houle, D. (1996). The lek paradox, condition dependence and genetic variance in sexually selected traits. *Proceedings of the Royal Society of London B*, 263, 1415–1421.
- Rushton, J. P. (2004). Placing intelligence into an evolutionary framework or how g fits into the r-K matrix of life history traits including longevity. *Intelligence*, 32, 321–328.
- Rushton, J. P. (2000). *Race, evolution, and behavior: A life-history perspective (3rd Edition)*. Port Huron, MI: Charles Darwin Research Institute.
- Rushton, J. P. (1985). Differential K theory: The sociobiology of individual and group differences. *Personality & Individual Differences*, 6(4), 441–452.
- SAS Institute, Inc. (1999). *SAS/STAT User's Guide, Version 8, Volumes 1, 2, and 3*. Cary, NC: SAS Institute.
- Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society of London. Series B*, 266, 1913–1917.
- Sefcek, J. A., Brumbach, B. H., Vásquez, G., & Miller, G. F. (2006). The Evolutionary Psychology of Human Mate Choice: How Ecology, Genes, Fertility, and Fashion Influence our Mating Behavior. Kauth, M. R., (Ed.). *Handbook of the Evolution of Human Sexuality Part 1 [Special Issue]*. *Journal of Psychology & Human Sexuality*, 18(2/3).
- Sefcek, J. A., & King, J. E. (2007) Chimpanzee Facial Symmetry: A Biometric Measure of Chimpanzee Health. *The American Journal of Primatology*. 69, 1–7.

- Sefcek, J. A., Miller, G. F., & Figueredo, A. J. (2007). Development and of an 18-item short form of the Ravens Advanced Progressive Matrices (RAPM-18). (Submitted).
- Segerstråle, U. (2000). *Defenders of the truth: The sociobiology debate*. New York: Oxford University Press.
- Sivacek, J., & Crano, W. D. (1982). Vested interest as a moderator of attitude-behavior consistency. *Journal of Personality and Social Psychology*, 43, 210-221.
- Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Taylor, C. E., & Condra, C. (1980). R- and K-selection in *Drosophila pseudoobscura*. *Evolution*, 34, 1183-1193.
- Thornhill, R. & Gangestad, S. W. (1999). The scent of symmetry: A human sex pheromone that signals fitness? *Evolution & Human Behavior*, 20(3), 175-201.
- Vernon, P. E. (1933). Some characteristics of the good judge of personality. *Journal of Social Psychology*, 4, 42-57.
- Whalley, L. J., Deary, I. J. (2001). Longitudinal cohort study of childhood IQ and survival up to age 76. *British Medical Journal*, 322, 819-822.
- Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-Item Short-Form Health Care Survey (SF-36): I. Conceptual Framework and Item Selection. *Medical Care*, 30, 473-483.
- Weibe, R. P. (2004). Psychopathy and sexual coercion: A Darwinian analysis. *Counseling and Clinical Psychology Journal*, 1(1), 24-41.
- Wicker, A. W. (1969). Attitudes vs. Action: The relationship of verbal and over behavioral responses to attitude objects. *Journal of Social Issues*, 25, 41-78.
- Wickett, J. C., Vernon, P. A., & Lee, D. H. (1994). In vivo brain size, head perimeter, and intelligence in a sample of healthy adult females. *Personality and Individual Differences*, 16, 831-838.
- Williams, G. C. (1957). Pleiotropy, natural selection, and the evolution of senescence. *Evolution* 11, 398-411.
- Williams, G. C., & Nesse, R. M. (1991). The dawn of Darwinian medicine. *Quarterly Review of Biology*, 66, 1-22.

- Williams, P. D., & Day, T. (2003). Antagonistic Pleiotropy, Mortality Source Interactions, and the Evolutionary Theory of Senescence. *Evolution*, 57(7), 1478-1488.
- Wilson, M., & Daly, M. (1985). Competitiveness, risk-taking, and violence: The young male syndrome. *Ethology and Sociobiology*, 6, 59-73.
- Zachary, R. A. (1986). *Shipley Institute of Living Scale, Revised Manual*. Los Angeles, CA: Western Psychological Services.
- Zahavi, A. (1975). Mate selection—a selection for a handicap. *Journal of Theoretical Biology*, 53, 205-214.
- Zahavi, A. (1991). On the definition of sexual selection, Fisher's model, and the evolution of waste and of signals in general. *Animal Behavior*, 42, 501-503.
- Zebrowitz, L. A., Hall, J. A., Murphy, N. A., & Rhodes, G. (2002). Looking smart and looking good: Facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin*, 28(2), 238–249.

TABLES

Table 1. Descriptors of traits associated with r-selection and K-selection

<u><i>r</i>-selected</u>	<u><i>K</i>-selected</u>
Smaller size	Larger size
Rapid Sexual Development	Slow Sexual Development
High Fertility (litters)	Low Fertility (singletons)
High Infant Mortality	Low Infant Mortality
Low Parental Investment	High Parental Investment
Short Lived	Long Lived
Variable Population Size	Constant Population Size
Reduced Competition	High Competition
Unstable Environments	Stable Environments

Table 2. Sample size comparisons and Cronbach's alpha's across measures

	<i>N</i>	Cronbach's α
<i>Measures of Somatic Investment</i>		
<i>Arizona K-Battery</i>		
Mini-K	172	.71
Parental Investment	190	.86
Family Support	191	.89
Friend Support	191	.84
Altruism	189	.79
Romantic Attachment	188	.88
Psychopathic Personality Inventory	182	.76
Risk Taking Questionnaire	191	.89
Mating Effort Scale	191	.70
<i>Measures of Neuro-developmental Stability</i>		
<i>General Intelligence (g)</i>		
Advanced Progressive Matrices (18-item)	194	.71
Mill-Hill	183	.88
Shipley Abstractions	193	.89
Shipley Vocabulary	193	.91
GPA	177	n/a
SAT	145	n/a
<i>Mental Health</i>		
Subjective Well-Being	188	.79
Beck Depression Inventory	190	.92
Hopkins Depression Index	188	.90
Hopkins Anxiety Index	188	.81
<i>Physical Health</i>		
SF-36	192	.89
General Health	192	.73
Symptoms (Brim et al. 2000)	190	.59

Table 3a. Factor loadings for two-factor solution of *K-Factor*

	<i>K-Factor</i> (44%)	<i>P-Factor</i> (41%)
Mini- <i>K</i>	.64	
Parental Investment	.57	
Family Support	.63	
Friend Support	.54	
Altruism	.25	
Romantic Attachment	.35	
Psychopathic Personality Inventory		.89
Risk Taking Questionnaire		.79
Mating Effort Scale		.70

Table 3b. Factor loadings for single-factor solution of *F-Factor*

	<i>F-Factor (76%)</i>
General Health	.47
SF-36	.73
SWB	.70
Negative Symptoms	-.60
Hopkins Anxiety	-.75
Hopkins Depression	-.88
BDI	-.84
Body Mass Index	.20
Facial Fluctuating Asymmetry	-.03
Body Fluctuating Asymmetry	-.04

Table 3c. Factor loadings for single-factor solution of *g-factor*

	<i>g-factor</i> (70%)
APM-18	.64
Mill Hill Vocabulary	.57
Shipley Vocabulary	.63
Shipley Abstractions	.54
GPA	.25
SAT	.35

Table 4. Higher-order factor loadings (reported as standardized regression coefficients)

	<i>Factor Loading</i>
<i>K-Factor</i>	
Mini-K	.70
Parental Investment	.63
Family Support	.68
Friend Support	.59
Altruism	.59
Romantic Attachment	.39
<i>P-Factor</i>	
Psychopathic Personality Inventory	-.61
Risk Taking Questionnaire	-.43
Mating Effort Scale	-.41
<i>F-Factor</i>	
Subjective Well-Being	.74
Beck Depression Inventory	-.85
Hopkins Depression Index	-.88
Hopkins Anxiety Index	-.80
SF-36	.79
General Health	.58
Medical Symptoms	-.48
<i>g-factor</i>	
APM-18	.63
Mill-Hill	.67
Shipley Abstractions	.70
Shipley Vocabulary	.57
GPA	.52
SAT	.66

Table 5a. Correlations among higher order factors

	<i>K-Factor</i>	<i>P-Factor</i>	<i>F-Factor</i>	<i>g-factor</i>
<i>K-Factor</i>	1.00			
<i>P-Factor</i>	-.01	1.00		
<i>F-Factor</i>	.31*	.12	1.00	
<i>g-factor</i>	-.06	-.14	-.06	1.00
* $p < .001$				

Table 5b. Correlations among higher order factors, by sex

Females ($n = 100$)				
	<i>K-Factor</i>	<i>P-Factor</i>	<i>F-Factor</i>	<i>g-factor</i>
<i>K-Factor</i>	1.00			
<i>P-Factor</i>	.03	1.00		
<i>F-Factor</i>	.45*	.18	1.00	
<i>g-factor</i>	-.07	-.11	-.05	1.00
Males ($n = 94$)				
	<i>K-Factor</i>	<i>P-Factor</i>	<i>F-Factor</i>	<i>g-factor</i>
<i>K-Factor</i>	1.00			
<i>P-Factor</i>	.00	1.00		
<i>F-Factor</i>	.14	.02	1.00	
<i>g-factor</i>	-.06	-.18	-.07	1.00
* $p < .001$				

Table 6. Correlations of *K* and *F* with particular 'g' measures

	<i>K-Factor</i>	<i>P-Factor</i>	<i>F-Factor</i>
APM-18	-.09	-.04	-.11
Mill-Hill	-.06	-.20**	-.06
Shipley Abstractions	-.06	-.02	-.07
Shipley Vocabulary	-.10	-.14*	-.02
GPA	.15*	-.20**	.17*
SA	.02	.11	-.01
* $p < .05$			

Table 7. Multiple Regression Statistics for the Theoretical Models

Predictor	Criterion Variable											
	<i>K-Factor</i>			<i>F-Factor</i>			<i>P-Factor</i>			<i>g-factor</i>		
	β weight	<i>t</i> test	<i>p</i> ($\beta > 0$)	β weight	<i>t</i> test	<i>p</i> ($\beta > 0$)	β weight	<i>t</i> test	<i>p</i> ($\beta > 0$)	β weight	<i>t</i> test	<i>p</i> ($\beta > 0$)
<i>K-Factor</i>				0.44	4.90	.001	0.02	0.32	.751	-0.06	-0.61	.545
sex				0.23	1.71	.089	0.56	5.14	.001	-0.01	-0.05	.962
<i>K-sex</i>				-0.20	-2.18	.030	-0.02	-0.22	.823	-0.01	-0.08	.937
<i>F-Factor</i>	0.46	5.02	.001				0.13	1.75	.081	-0.04	-0.42	.674
sex	-0.19	-1.36	.175				0.54	5.01	.001	0.00	0.01	.989
<i>F-sex</i>	-0.22	-2.40	.017				-0.08	-1.08	.281	-0.03	-0.27	.787
<i>K-Factor</i>										-0.05	-0.45	.651
<i>F-Factor</i>										0.02	0.19	.852
<i>KF</i>										0.09	0.89	.376
sex										0.02	0.15	.877
<i>K-sex</i>										-0.03	-0.26	.795
<i>F-sex</i>										-0.09	-0.85	.397
<i>KF-sex</i>										0.05	0.59	.558

FIGURES

Figure 1. Example of the interaction between the G-matrix and M-matrix

Figure 2. Differential-K model and general intelligence

Figure 3. Fitness factor model and general intelligence

Figure 4. Hybrid model of the relationships among K , F , and g

Figure 5. Points used in calculating facial asymmetry measurements

APPENDIX A

CHIMPANZEE FACIAL SYMMETRY: A BIOMETRIC MEASURE OF

CHIMPANZEE HEALTH

Abstract

This paper reports a study of fitness indicator theory in chimpanzees. First, it establishes a theoretical perspective for the study of fitness indicator theory and the relationships among indicators of fitness in humans and other animals. Second, it describes a methodology for assessing facial fluctuating asymmetry (FA) in a sample ($N = 21$) of zoo chimpanzees (*Pan troglodytes*). Third, associations among chimpanzee facial FA and health are described. FA was positively associated with negative health symptoms, and negatively associated with general health. Results are discussed under the framework of *good genes* theory.

Introduction

Sexual selection provides a theoretically important perspective on health and behavior within both the social and biological sciences. One subset of sexual selection, *fitness indicator theory*, suggests that many phenotypic traits are external displays of underlying genetic quality i.e. lower deleterious mutation loads (Miller, 2000). These fitness indicators are affected by environmental variables including nutrition, parasitism, physical traumas, and disease, making them phenotypic markers of developmental stability in response to environmental contaminants. Fundamentally, fitness indicator theory is about costly signaling, wherein traits that seem wasteful or arbitrary are shaped by selection because they convey reliable, honest information about an individual's resistance to environmental perturbations. In turn, these signals work to benefit both the signaler and the receiver, by increasing the reproductive fitness of those who produce high-quality signals, and those who can discern the signals and use them accurately in mate choice (Zahavi, 1975, 1991).

For example, in humans facial and body fluctuating asymmetry (FA) (deviations from perfect bilateral symmetry) appears to be related to underlying fitness-related traits including facial masculinity and physical robustness (Gangestad & Thornhill, 2003; Manning & Pickup, 1998), and susceptibility to various physical (see Moller & Swaddle, 1997, for review) and mental health problems (Martin, Manning, & Dowrick, 1999). Further, lower FA seems to be related to more positive ratings on psychological traits (e.g. personality, intelligence, subjective well-being), as well as increased number of sexual partners over the lifetime (see Sefcek, Brumbach, Vasquez, & Miller, 2006, for

review). It is believed that these asymmetries begin developing in the fetal environment as the result of either genetic or environmental stress affecting a normally well-canalized ontogenic pathway (Kohn & Bennet, 1986).

In non-human primates, dental and dermatoglyphic asymmetries have been shown to increase with prenatal maternal stress in pigtailed (Newell-Morris, Fahrenbruch, & Sackett, 1989) and rhesus macaques (Kohn & Bennet, 1986); but this proposition is not without criticism (e.g. see Saunders & Mayhall, 1982). Other studies, have suggested that the stress hypothesis extends beyond the fetal environment to encompass all aspects of ontogenetic development (Bailit, Workman, Niswander, & MacLean, 1970). Researchers including Manning and Chamberlain (1993; 1994), have identified significant relationships between canine asymmetry and several measures of sexual selection (e.g. canine dimorphism, canine size, and intrasexual competition), suggesting that such asymmetries are the result of directional selection on traits signaling either genetic quality, the ability to buffer environmental perturbations, or both.

Jacobson, Steklis, Steklis, LeClair, and Fawcett (2004) have shown that facial FA could be reliably measured from digital photographs of wild mountain gorillas, but such measures have not yet been related to any psychological, health, or ecological variables. Although there are studies in non-human species concerning health outcomes as related to symmetry (Robins & Rogers, 2002), there is, to our knowledge, no study incorporating an external measure of facial symmetry. The current study therefore had two main goals. First, it was designed to develop a methodology for measuring facial FA in a population of zoo-housed common chimpanzees. Second, it examined the association between

symmetry and health in chimpanzees.

Method

Subjects

Health ratings were collected from 30 female and 19 male chimpanzees (*Pan troglodytes*) housed at 13 zoological parks participating in the ChimpanZoo program of the Jane Goodall Institute. The mean age of these chimpanzees was 29.98 years (male = 28.73, female = 30.74) with a $SD = 13.71$ (male = 10.83, female = 15.46). Due to the difficulty in obtaining useable facial photographs (see below), fluctuating asymmetry analyses were conducted on a subsample of 12 female and 9 male chimpanzees with a mean age of 33.91 years (male = 30.33, female = 36.29) and $SD = 16.59$ (male = 17.39, female = 16.62).

Raters

Health ratings were collected by zoo employees. Facial asymmetry analyses were made by trained research assistants. Each chimpanzee was rated by a mean of 4.05 raters for the health measures, and 2 raters for the facial asymmetry analysis.

Materials

Chimpanzee General Health Scale.

This scale was adapted from a human health scale used by Figueredo et al. (2006). The chimpanzee version of the scale contained five items asking zoo employees to rate the overall physical and mental health of the chimpanzee (two at the time of the ratings, two for when the chimpanzees were adolescents, and one in comparison to the rest of the chimpanzee group). Questions 1-4 were scored on a 0-5 scale coded as: (0) poor, (1) fair,

(2) good, (3) very good, (4) excellent, (5) don't know/not applicable. Because question five was asking for a comparison to other chimpanzees it was scored on a 0-4 scale as follows; (0) much worse, (1) somewhat worse, (2) about the same, (3) somewhat better, and (4) much better. A general health score was calculated by the average of the summed numerical value of all responses, so that higher scores indicated better health. Items are listed below:

1. In general, would you say the chimpanzee's *current* PHYSICAL HEALTH is...
2. In general, would you say the chimpanzee's *current* MENTAL HEALTH is...
3. How was this chimpanzee's PHYSICAL HEALTH *when an adolescent* (ages 4 to 12)?
(If the chimpanzee is either too young or you do not know please indicate a '5')
4. How was this chimpanzee's MENTAL HEALTH *when an adolescent* (ages 4 to 12)?
(If the chimpanzee is either too young or you do not know please indicate a '5')
5. In general, compared to the other chimpanzees in the group, how would you say this chimpanzee's *overall* health is?

Chimpanzee Medical Symptoms Scale.

Medical Symptoms were estimated by the endorsement of 27 possible medical conditions. Items for this scale were taken from a 29-item medical symptoms scale used by Figueredo et al. (2006), on a human sample. To determine the suitability of the items on the scale, we asked several primate experts for their opinion on the suitability of the listed medical conditions, such as "Thyroid disease," "High blood pressure or hypertension," or "Ulcer." Two items were dropped from the original scale, namely, "Alcohol or drug problems" and "Migraine Headaches". We used the sum of all endorsed items as a measure of negative symptomology.

Facial Fluctuating Asymmetry Measurements

We used a mixture of techniques previously used for measuring facial asymmetry

of gorillas (Jacobson, et al. 2004) and humans (Grammer & Thornhill, 1994). Horizontal FA (HFA) measurements, the degree to which left/right side traits deviate from each other along a horizontal plane, were calculated by taking the relative differences between bisected lines from several facial landmarks. We created a midline defined by the average midpoints of eight bisected lines (described below), then subtracted the midpoint of each bisected line from this average midline and took its absolute value.

Because we relied on our own methodology, facial markers were chosen over the course of several months of practice measuring proposed points. Retained points fit two criteria. First, they were facial landmarks that were unobstructed by facial hair across the sample. Second, they showed relatively small individual differences across individuals, i.e. they *looked* species-typical, and measurable across the sample. Eight out of 12 proposed points were retained.

These 8 paired locations included: D1: Distance between lower protrusions of brow ridge; D2: Distance between centers of pupils; D3: Distance between inner eyes; D4: Distance between outer eyes; D5: Distance between upper nostrils; D6: Distance between widest protrusions of nostrils; D7: Width of the collumna between nostrils; and D8: Distance between lower nostrils (see Figure 1). The absolute values of these right-left differences were then summed to define a facial HFA composite score. Vertical FA (VFA) measurements were also taken by summing the differences between the vertically bisected lines of the same eight measures. Each FA composite score was then standardized and the components summed for a complete FA composite score (CFA).

Results

Analyses

FA measurements were conducted using Image-J version 1.31 software (Rasband, 2003), then imported into SAS (SAS Institute, 1985) for further analysis.

Health Scale Reliabilities

Internal consistencies (*IC*) were estimated with Cronbach's alpha. In addition, two measures of interrater reliability were calculated. $ICC(3, 1)$ estimated the reliability of individual raters, where: $ICC(3, 1) = [MS_{(chimps)} - MS_{(items \times chimps)}] / [MS_{(chimps)} + [(k-1) \times MS_{(items \times chimps)}]$. $ICC(3, k)$ estimated the reliability of mean scores based on a mean of 4.05 raters for each chimpanzee health measure, and 2 raters for the FA measures, where: $ICC(3, k) = [MS_{(chimps)} - MS_{(items \times chimps)}] / MS_{(chimps)}$. The mean square for the Raters x Chimpanzees interaction was used as the error term for both reliability estimates.

Internal consistency for the general health scale was moderate ($\alpha = .72$), with moderate interrater reliabilities ($ICC(3, 1) = .49$; and $ICC(3, k) = .24$). Internal consistency for the symptoms measure was low ($\alpha = .38$), with high interrater reliability, ($ICC(3, 1) = .90$; and $ICC(3, k) = .95$).

Fluctuating Asymmetry Reliabilities

Two raters were used to assess asymmetry reliability. Interrater reliabilities for individual measurements and composite scores are listed in Table 1. Across all measures, internal consistencies of the composite scores were moderate to strong (for all $\alpha = .72$ to .91). However, for individual FA measurements ($n = 16$) only half showed $ICC(3, k)$ values greater than .50. Therefore, further FA analyses defined composite horizontal FA

(HFA), vertical FA (VFA), and total composite FA (CFA) with the eight measurements that had $ICC(3,k)$ values greater than .50. Based on these eight items, intertrater reliabilities were recalculated for the new composite scores. These reliabilities remained moderate to strong (for all $\alpha \geq .65$ to .88).

Relationship between FA and Health

Type I hierarchical General Linear Models (PROC GLM) were used to control for the relationships between age and general health, and age and negative health symptoms ($r = -.36, p = .031$; and $r = .54, p < .001$, respectively). The model defined each of the health measures as criteria with the ordered predictors being age, asymmetry measurement, and age x asymmetry measurement. For general health, VFA was negatively related to health after age related variance was removed ($F = 5.02, p = .049$). A similar but nonsignificant trend was noted for CFA ($F = 4.14, p = .069$). Both CFA and VFA were positively related to negative symptoms after age effects were removed ($F = 17.53, p = .002$; and $F = 11.65, p = .007$).

Discussion

Good genes theory states that FA signals are valid advertisements of genetic quality by which others can gauge the fitness of potential mates and allies (see Sefcek et al., 2006). The purpose of the current study was to explore the potential of using chimpanzee facial asymmetry as a biometric measure of health.

Overall composite measures of chimpanzee facial FA predicted general health and negative symptoms. Specifically, high FA was positively associated with zoo keeper reports of negative health symptoms, and negatively associated with reports of lower

overall health. These findings are consistent with studies in relation to human mate-choice preferences for facial symmetry and its relationship to physical and mental health (see Sefcek et al., 2006, for review).

In many species FA, also appears to be related to a variety of fitness effects. For example, numerous studies indicate a negative relationship between FA and the quantity and quality of sperm in antelopes (Gomendio et al., 2000), insects (Farmer & Barnard, 2000), and humans (Manning, Scutt, & Lewis-Jones, 1998). Many other studies show female preferences for low FA in, for example, wings of male scorpion flies (Thornhill, 1992); tails of barn swallows (Møller, 1992); and eyespots of peacock tails (Petrie, Halliday & Sanders, 1991). Studies that explore the effect of parental FA on offspring condition (infant mortality, parasite resistance, fertility, and fecundity), have shown a similar pattern of results, with FA being negatively related to offspring condition (see Swaddle, 2003, for review).

The current research suggests that chimpanzee facial FA may be a useful metric to gauge chimpanzees' overall health and condition. Because there are some sex differences in correlates of FA in humans (Sefcek, et al., 2006), sex differences in chimpanzee FA-health relationships should also be studied. Further, if indeed FA is a marker of overall developmental stability then the *good genes* model suggests there should be correlations among all aspects of phenotype that tap into this genetic architecture. Aspects of chimpanzee personality, intelligence, social behavior, and health should all show some relationship to a biological marker of developmental stability including FA.

Acknowledgements

We are indebted to Dr. V. Landau, Director of ChimpanZoo, and Erica Metelovski, Assistant to the Director of ChimpanZoo, for their aid in data collection, and our research assistant Daniel B. Galloway. We would also like to thank Eric Matthews and Robert Cooper for their donations of many of the digital photographs used in this study, Dr. H.D. Steklis for his methodological guidance, as well as the following facilities and their many personnel for providing chimpanzee ratings: Cheyenne Mountain Zoo (Colorado Springs), Dallas Zoo, Lion Country Safari (West Palm Beach), Los Angeles Zoo, Lowry Park Zoological Gardens (Tampa Bay), North Carolina Zoological Park (Asheboro), Oakland Zoo, Sacramento Zoo, San Francisco Zoo, Sedgwick County Zoo (Wichita), Sunset Zoo (Manhattan, Kansas), Taronga Zoo (Mossman, Australia), and Tulsa Zoo. Portions of these data were previously reported at the ChimpanZoo Conference, Salina, KS, October, 2005.

References

- Bailit, H. L., Workman, P. L., Niswander, J. D., & MacLean, C. J. (1970). Dental asymmetry as an indicator of genetic and environmental conditions in human populations. *Human Biology*, 42, 626 - 638.
- Farmer, D. C., & Barnard, C. J. (2000). Fluctuating asymmetry and sperm transfer in male decorated field crickets (*Gryllodes sigillatus*). *Behavioral Ecology and Sociobiology* 47, 287 - 292.
- Figueredo, A. J., Vasques, G., Brumbach, B. H., & Schneider, S. M. R. (2006). The heritability of life history strategy: The *K-Factor*, covitality, and personality. *Social Biology*, 51, 121 - 143.
- Gangestad, S. W., & Thornhill, R. (2003). Facial masculinity and fluctuating asymmetry. *Evolution and Human Behavior* 24, 231 - 241.
- Gomendio, M., Cassinello, J., & Roldan, E. R. S. (2000). A comparative study of

- ejaculate traits in three endangered ungulates with different levels of inbreeding: Fluctuating asymmetry as an indicator of reproductive and genetic stress. *Proceedings of the Royal Society of London B*, 267, 875-882.
- Grammer, K., & Thornhill, R. (1994). Human (*Homo sapiens*) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology*, 108, 233 - 242.
- Jacobson, A., Steklis, H. D., Steklis, N. G., LeClair, J., & Fawcett, J. K. (2004). Exploring the application of human facial fluctuating asymmetry methodology in a great ape. Paper. In Steklis, H. D. *New Approaches to Facial Attractiveness* (Chair). Symposium. Annual Meeting of the Human Behavior and Evolution Society, Berlin, Germany.
- Kohn, L. A. P., & Bennett, K. A. (1986). Fluctuating asymmetry in fetuses of diabetic rhesus macaques. *American Journal of Physical Anthropology*, 71, 477 - 483.
- Manning, J. T., & Chamberlain, A. T. (1993). Fluctuating asymmetry, sexual selection, and canine teeth in primates. *Proceedings of the Royal Society of London B*, 251, 83 - 87.
- Manning, J. T., & Chamberlain, A. T. (1994). Fluctuating asymmetry in gorilla canines: a sensitive indicator of environmental stress. *Proceedings of the Royal Society of London B*, 255, 189 - 193.
- Manning, J. T., & Pickup, L. J. (1998). Symmetry and performance in middle distance runners. *International Journal of Sports Medicine*, 19, 1 - 5.
- Martin, S. M., Manning, J. T., & Dowrick, C. F. (1999). Fluctuating asymmetry, relative digit length, and depression in men. *Evolution and Human Behavior*, 20, 203 - 214.
- Manning, J. T., Scutt, D., & Lewis-Jones, D. I. (1998). Developmental stability, ejaculate size and sperm quality in men. *Evolution and Human Behavior*, 19, 273 - 282.
- Miller, G. F. (2000). Mental traits as fitness indicators: Expanding evolutionary psychology's adaptationism. *Annals of the NY Academy of Science*, 907, 62 - 74.
- Møller, A. P. (1992). Female swallow preference for symmetrical male sexual ornaments. *Nature*, 357, 238 - 240.
- Møller, A. P., & Swaddle, J. P. (1997). *Asymmetry, developmental stability, and evolution*. Oxford University Press, Oxford, UK.

- Newell-Morris, L. L., Fahrenbruch, C. E., & Sackett, G. P. (1989). Prenatal psychological stress, dermatoglyphic asymmetry and pregnancy outcome in the pigtailed macaque (*Macaca nemestrina*). *Biology of the Neonate*, 56, 61 - 75.
- Petrie, M., Halliday, T., & Sanders, C. (1991). Peahens prefer peacocks with elaborate trains. *Animal Behavior*, 41, 323 - 331.
- Rasband, W. (2003). Image J 1.31v. *The National Institute of Health*, <http://rsb.info.nih.gov/ij/>.
- Robins, A., & Rogers, L. J. (2002). Limb preference and skeletal asymmetry in the cane toad, *Bufo marinus* (Anura: Bufonidae). *Laterality: Asymmetries of Body, Brain, and Cognition* 7(3): 261 - 275.
- SAS Institute Inc. (1985). *SAS user's guide: Basics Version 5*. NC: SAS Institute.
- Saunders, S., & Mayhall, J. (1982). Fluctuating asymmetry of dental morphological traits: new interpretations. *Human Biology*, 54, 789 - 799.
- Sefcek, J. A., Brumbach, B. H., Vásquez, G., & Miller, G. F. (2006). The Evolutionary Psychology of Human Mate Choice: How Ecology, Genes, Fertility, and Fashion Influence our Mating Behavior. Kauth, M. R., (Ed.). Handbook of the Evolution of Human Sexuality Part 1 [Special Issue]. *Journal of Psychology and Human Sexuality*, 18(2/3), 125 - 182.
- Swaddle, J. P. (2003). Fluctuating asymmetry, animal behavior, and evolution. *Advances in the Study of Behavior*, 32, 169 - 205.
- Thornhill, R. (1992). Fluctuating asymmetry and the mating system of the Japanese scorpionfly, *Panorpa japonica*. *Animal Behavior*, 44, 867 - 879.
- Zahavi, A. (1975). Mate selection—a selection for a handicap. *Journal of Theoretical Biology*, 53, 205 - 214.
- Zahavi, A. (1991). On the definition of sexual selection, Fisher's model, and the evolution of waste and of signals in general. *Animal Behavior*, 42, 501 - 503.

Table 1
Reliabilities for facial asymmetry measurements

	FA Measurement								
	HFA	HD1	HD2	HD3	HD4*	HD5*	HD6*	HD7	HD8*
IC	.91 (.88*)								
ICC (3,k)	0.74 (0.93*)	0.16	0.2	0.23	0.67	0.89	0.82	0.33	0.66
ICC (3,1)	0.58 (0.86*)	0.09	0.11	0.13	0.26	0.79	0.70	0.20	0.50
	VFA	VD1*	VD2	VD3*	VD4	VD5*	VD6	VD7*	VD8
IC	.72 (65*)								
ICC (3,k)	0.58 (0.85*)	0.53	-0.71	0.69	-1.65	0.90	0.46	0.88	0.46
ICC (3,1)	0.40 (0.74*)	0.36	-0.26	0.53	-0.45	0.81	0.31	0.78	0.30
	CFA								
IC	.90 (.87*)								
ICC (3,k)	0.89 (0.93*)								
ICC (3,1)	0.80 (0.87*)								

* Indicates measurements retained for further analysis

Figure 1: Measurement Points for Human and Chimpanzee Facial Asymmetry

APPENDIX B

DEVELOPMENT AND VALIDATION OF AN 18-ITEM MEDIUM FORM OF THE
RAVENS ADVANCED PROGRESSIVE MATRICES (APM-18)

Abstract

The Ravens Advanced Progressive Matrices (APM) is a widely used measure of general intelligence (*g*), both across settings and cultures. Due to its lengthy 40-minute administration time, several researchers have developed short form scales, yet these forms typically yield a significantly lower reliability. This paper describes the creation of an 18-item short form (APM-18), and its validation in three samples of Southwestern U.S. university students (total *N* = 633). The APM-18 shows similar psychometric properties to both the previously published 36-item long form and 12-item short form, but retains a reliability estimate closer to the original APM. This, plus the shorter administration time (25 minutes) relative to the complete APM (40-60 minutes), makes it useful for time-constrained or mass testing situations.

Introduction

The Ravens Progressive Matrices Test, developed by Raven (1943) as a measure of general intelligence (g), has undergone many revisions, ranging from colored versions for children to the Standard and Advanced Matrices for adults of different cognitive levels. The most recent published version is the Ravens Advanced Progressive Matrices (APM) (Raven, Raven, & Court, 1993), which was developed for higher-ability adult populations (i.e. college level and above). This test is constructed of 36-items of increasing difficulty broken into three 12-item sets; in each item, the examinee is asked to complete a visual pattern by choosing 1 of 8 possible solutions.

Due to its nonverbal format, the APM is purported to be a culturally-fair, unbiased measure of fluid intelligence (Cattell, 1963), eductive ability (Raven et al., 1993), or, as we will refer to it, general intelligence (g ; Spearman, 1927), and has shown itself to be especially useful in situations where English is not an individual's primary language. As such, the Standard and Advanced Progressive Matrices have been used extensively in many applied settings in the U.S. (e.g. Ackerman, 1992) and across many cultures (Owen, 1992; Raven, 2000; Rushton, Cvorovic, & Bons, 2007). However, the positive aspects of this test are marred by its lengthy administration time (40-60 minutes), making it difficult to use in time-constrained multivariate research or classroom settings.

In answer to these various limitations, Arthur and Day (1994) developed a 12-item short form of the APM (which we call APM-12), with an administration time of 15 minutes. Several studies have shown that this 12-item form shows acceptable psychometric properties (e.g. Cronbach's alpha, test-retest reliability, convergent

validity) (see Arthur, Tubre, Paul, & Sanchez-Ku, 1999, for review). However, this short form shows relatively low and variable internal consistency. For example, Cronbach's alphas range from .58 - .66 for short-form itself to and .72 - .73 for the 12 short-form items extracted from the full 36-item version (Arthur et al., 1992).

More recently, Hamel and Schmittmann (2006) have argued that the complete 36-item APM can be administered as a 20-minute *speed test*. Scores on this speeded form of the APM shows strong correlations with scores on slower timed (40 minutes, $r = .74$) and untimed versions ($r = .75$) of the APM. However, these authors failed to report the internal consistency of the *speed test* scale. We also suspect that giving typical adults only 20 minutes to complete 36 very challenging abstract reasoning problems might impose undue stress.

The purpose of the current study was to develop a medium-form version of the APM that resulted in higher internal consistency than the 12-item version (APM-12), but shorter administration time than the full 36-item APM (APM-36) – a combination of features that might be useful for time-constrained and mass testing situations. Here we report the development and construct validity of this 18-item scale.

Study 1: Scale Construction and Construct Validity

Method

Participants

A total of 633 students (198 male, 435 female) from three Southwestern Universities participated in this study as a partial requirement for experimental course credit. The mean age for participants was 20.92, $SD = 4.07$ ($M_{male} = 20.85$, $SD_{male} = 3.90$;

$M_{female} = 20.96$, $SD_{female} = 4.15$). Ages ranged from 17 to 58 years old (male = 18 to 41; female = 17 to 58).

Measures

The Ravens Advanced Progressive Matrices 18-item Short-Form (APM-18)

This 18-item short-form version of the APM is printed in a booklet format on 8 ½ by 11" white paper, with each test-item printed on a separate page. The first four pages of the test booklet contain three example items (practice items 1, 5, and 9 from APM-36) to explain the task.

The 18 actual test-items were derived by adding 6 items from the longer 36-item version (Raven et al. 1993) to Arthur and Day's (1994) published 12-item version. Arthur and Day used items 1, 4, 8, 11, 15, 18, 21, 23, 25, 30, 31, and 35 from the 36-item APM based on a set of three decision rules which can be summed up as 1) dividing the APM into 12, 3-item sections based on difficulty, 2) taking the item with the highest item-total correlation for each section, and 3) in the case of a tie, including the item that resulted in the largest drop in internal consistency if it was excluded from the full test. Following these same rules, we added 6 more items of increasing difficulty – two that were easy (96 and 75% of examinees from the normative sample answered correctly), two that were moderate (50 and 48% of examinees from the normative sample answered correctly), and two that were difficult (37 and 32% of examinees from the normative sample answered correctly). These items (2, 20, 22, 24, 34, and 32) were integrated in order of difficulty to mimic their presentation order in the original APM.

Procedure

The new APM-18 test was given in classroom settings with several examinees at a time. This was done because this test was developed as a measure of g that could be used in environments such as classrooms, where there are time limits on research sessions. In one subsample ($n = 175$), tests were given with no time constraints, but with completion times recorded, to determine the average time needed for completion. The other two subsamples ($n = 232$ and $n = 226$) were constrained to finish the test within 25 minutes, with no individual completion times recorded.

Analyses

All statistical analyses were conducted using SAS version 8.2 (SAS Institute, 1999). Cronbach's alphas and bivariate correlations were computed using the *PROC CORR* procedure. Tests for mean differences between sexes were calculated through *t*-test (*PROC TTEST*) procedures. Hierarchical General Linear Models were tested using *PROC GLM*.

Results

Internal consistency (IC) estimates were computed by using Cronbach's alpha. The IC of the APM-18 scale yielded moderate reliability ($\alpha = .79$). This alpha is lower than normative IC reports for the APM-36 ($\alpha = .84$; Forbes, 1964), but higher than those for the APM-12 (ranging from $\alpha = .58$ - $.66$; see Arthur et al. 1999). Further, the alpha of the APM-18 was larger than that of the embedded APM-12 ($\alpha = .73$). Table 1 shows the results for each of the APM-18 items, with respect to their item-total correlations, item difficulties, and scale α of the overall scale if the item is deleted.

The mean APM-18 score was 9.73, $SD = 3.59$ ($M_{male} = 10.43$, $SD_{male} = 3.52$;

$M_{female} = 9.41$, $SD_{female} = 3.59$), with a range of 18. For the subsample in which completion times were recorded ($n = 175$), the mean test completion time was 17.5 min ($SD = 4.67$), with a range of 7 to 25 minutes; 21% of participants took longer than 20 minutes, but no one took longer than 25 minutes. In this subsample, there was a significant positive relationship between the amount of time it took for participants to take the test and their APM-18 score ($r = .41$, $p < .001$), but there was no relationship between age and APM-18 score, or age and time required to complete the test (for each $r \leq .03$, $p \geq .71$). However, for the complete sample ($N = 633$), younger participants scored a little higher (age and APM-18 scores correlated $r = -.15$, $p < .001$), and males scored a little higher (sex [female = 0, male = 1] and APM-18 scores correlated $r = .13$, $p < .001$)

Hierarchical General Linear Models (GLMs) were tested to explore whether the apparent differences in male and female APM-18 scores might have been indirectly attributable to the relationship between age and APM-18 scores. This model defined the APM-18 score as the criterion variable, with the ordered predictor variables being age and then sex. The hierarchical model was designed to allow age to absorb as much variance as possible, with sex entered into the model only afterwards. Using this model, both GLMs indicated a significant effect for age ($F = 15.39$, $p < .001$) and then also for sex after age had been statistically controlled ($F = 10.66$, $p = .001$).

Discussion

The results presented here suggest that the APM-18 may serve as a useful compromise between the lower-reliability APM-12 and the much-longer APM-36. The hierarchical GLMs identify both age and sex to be significant predictors of APM-18

scores, with younger individuals and males generally scoring higher. These results are consistent with many previous studies looking at general intelligence (e.g. Jackson & Rushton, 2006). Results of Study 1, however, do not test the convergent validity of this scale relative to other measures of intelligence. Study 2 was designed to do this.

Study 2: Convergent Validity

Study 2 was conducted to assess the convergent validity of the APM-18 with other measures of intelligence, academic achievement, and personality. To do so we tested two separate sub-samples (n = 193 and 229) taken from the Study 1, each of which used different criterion measures. In Sample 1, two widely-used measures of adult intelligence were used; the Mill-Hill Vocabulary Scale – Multiple Choice Sets A & B (*MHV-MC*; Raven, Raven, & Court, 1997), developed to be used in conjunction with the APM-36 as a measure of *reproductive*, i.e. the ability to store and retrieve information (Raven, 1989); and the Shipley Institute of Living Scale (*SILS*; Zachary, 1986), which is a stand-alone intelligence test comprised of two subscales; Vocabulary, which tests crystallized intelligence; and Abstraction, which tests fluid intelligence. Also, we examined academic performance via self-reported GPA and SAT scores. In Sample 2, we examined correlations between APM-18 scores and Big Five personality dimensions assessed with the NEO-FFI scale (Costa & McCrae, 1992), and verbal and drawing creativity (Miller & Tal, 2007). Additionally, ACT scores were collected in this second sample.

Method

Participants

Sample 1 was comprised of 193 students (94 male, 99 female) from an introductory psychology course at the University of Arizona. Mean age of participants was 19.11, $SD = 1.62$ ($M_{male} = 19.23$, $SD_{male} = 1.07$; $M_{female} = 19.01$, $SD_{female} = 2.00$). Due to the length of time required to administer the APM-18, the Shipley, and the Mill-Hill, 10 participants did not complete the Mill Hill Test. We urged participants to record their SAT, ACT, and GPA scores only if they were certain of them; due to this constraint, many of these scores were also missing.

Sample 2 was comprised of 229 students (65 male, 164 female) from various undergraduate courses at the University of New Mexico. Mean age of participants was 20.19, $SD = 3.43$ ($M_{male} = 21.05$, $SD_{male} = 5.01$; $M_{female} = 19.85$, $SD_{female} = 2.48$). Again we urged participants to record their ACT scores only if they were certain of them, leaving us with ACT scores for only 129 participants.

Measures

APM-18. The APM-18 consisted of the same items identified in Study 1. In Sample 1 the form was presented first in a series of measures examining adult intelligence. In Sample 2 it was presented in the middle of a questionnaire packet concerning personality, creativity, sexual behavior, and intelligence.

The Shipley Institute of Living Scale (SILS) (Zachary, 1986). The SILS is a timed (10-minutes per subscale), 60-item self-report measure that examines both verbal intelligence (40 items) and abstract intelligence (20 items). The test is considered appropriate for average English-speaking individuals from 14 to adult ages, who are motivated test-takers. Validities and norms published in the manual were taken from a

sample of 322 army recruits. Split-half reliabilities for each subscale are reported as .87 for Vocabulary, .89 for Abstraction, and .92 for the total score.

The Mill-Hill Vocabulary Scale Senior Multiple Choice Form, Sets A & B (MHV-MC) (Raven, et al., 1997). The MHV-MC is a 68-item self-administered multiple-choice vocabulary test designed to complement the APM-36. Whereas the APM aimed to measure an individual's ability to solve novel problems and think in novel ways (i.e., *fluid* intelligence), the Mill-Hill aimed to measure an individual's ability to recall learned information (i.e., *crystallized* intelligence). To this extent it indicates educational attainments, cultural background, and familiarity with the test's language. The Mill-Hill typically shows split-half reliabilities over .90, and test-retest reliabilities ranging between .87 - .95 (Raven et al.).

Academic Performance. Academic performance was measured by self reported grade-point averages (GPA) and Scholastic Aptitude Test (SAT) scores in Sample 1. Sample 2 participants were asked for SAT and ACT scores. A variety of studies have identified moderate to strong correlations between these academic achievement and aptitude measures, and a variety of other traits, including intelligence, personality, and psychopathology (Barton, Dielman, & Cattell, 1971; Brown, 1994; Dyer, 1987; Mouw & Khanna, 1993)

NEO-Five Factor Inventory (NEO-FFI) (Costa & McCrae, 1992). The NEO-FFI is the most-widely used measure in research on the Five-Factor model of personality. It is a shortened version of the 240-item NEO-PI-R (Costa & McCrae, 1992), comprised of 60 items that measure five global personality factors (12 items per factor): Openness to

Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. In our version, participants rated degree of agreement with statements about their personalities and behavioral propensities on a five-point scale ranging from -2 (strongly disagree) to 0 (neutral) to +2 (strongly agree). This scale has shown strong internal consistency, with Cronbach's alphas ranging between .74 to .89, for each factor, and consistent cross-cultural validity (McCrae & Costa, 1997).

Verbal and Drawing Creativity Tasks (Miller and Tal, in press). Participants completed six 2-minute verbal creativity tasks and eight 1-minute drawing creativity tasks. Since a mating-oriented mind-set promotes creativity (Griskevicius, Cialdini, and Kenrick, 2006), participants were asked to complete these tasks as creatively as possible with the intention of attracting a romantic partner. Examples of verbal tasks included writing answers to thought-provoking questions, such as: "How would you keep a marriage exciting after the first couple of years?", "What do you hope the world will be like in a hundred years?", and "Imagine that all clouds had really long strings hanging from them – strings hundreds of feet long. What would be the implications of that fact for nature and society?" There were two types of drawing tasks, 4 abstract (e.g. "Please draw an abstract symbol, pattern, or composition that represents your happiness as a child doing a favorite activity"), and 4 representational (e.g. "In the space below, please draw an animal that you admire for its strength, grace, speed, or beauty"). Each participant's responses to each of the 14 creativity tasks was scored independently by four raters on a 1-5 creativity scale. The resulting composite verbal creativity and drawing creativity measures showed high inter-rater reliability and internal consistency (Cronbach's alphas

= .91 and .90, respectively) (Miller & Tal, in press).

Results

Sample 1

Internal consistency (IC) estimates were computed using Cronbach's alpha. The APM-18 showed moderate IC ($\alpha = .71$), with the embedded APM-12 yielding a slightly lower value ($\alpha = .63$). While these internal consistencies are lower than those reported in Study 1, they are still moderate in strength.

The mean APM-18 score was 10.68, $SD = 3.25$ ($M_{male} = 11.07$, $SD_{male} = 3.13$; $M_{female} = 10.31$, $SD_{female} = 3.34$), with a range of 13 (4 – 17). There was no relationship between APM-18 score and age ($r = .03$, $p = .71$), or sex ($r = .11$, $p = .10$). Mean scores for each of the measures of intelligence can be seen in Table 2. Due to the significant sex difference between APM-18 scores in Study 1, mean sex differences on all measures in this study were checked via t-tests. There were no significant sex differences for any of the intelligence measures in this sample, except for a moderate male advantage on self-reported SAT scores ($t = -3.00$, $p = .003$). Therefore, the remaining analyses were conducted on the full sample rather than by sex.

As seen in Table 3, both the APM-18 and embedded APM-12 correlated significantly with most of the other measures of intelligence and academic achievement and aptitude used in this sample.

Sample 2

Internal consistency (IC) estimates were again computed using Cronbach's alpha. As in Study 1 and Sample 1 of this study, the APM-18 showed moderate reliability ($\alpha =$

.79), while the embedded APM-12 again shows slightly lower reliability ($\alpha = .74$). The mean APM-18 score was 9.53, $SD = 3.57$ ($M_{male} = 10.29$, $SD_{male} = 3.98$; $M_{female} = 9.23$, $SD_{female} = 3.36$), with a range of 1 to 18. There was no relationship between APM-18 score and age ($r = -.03$, $p = .61$), but there was between APM-18 score and sex ($r = .13$, $p = .04$), with males again scoring slightly higher ($t = -2.04$, $p < .05$). There were no other sex differences on the other intelligence measures (for all $t \geq -1.20$, $p > .05$). Table 4 shows mean scores for the intelligence measures and NEO-FFI factors.

As seen in Table 5, both the APM-18 and embedded APM-12 scores were significantly positively related to verbal creativity ($r = .36$ and $.32$, $p < .001$, respectively), drawing creativity ($r = .34$ and $.29$, $p < .001$, respectively), and self-reported ACT score ($r = .44$ and $.45$, $p < .001$, respectively). Additionally, the APM scales were positively related to Openness ($r = .26$ and $.25$, $p < .01$, respectively) and negatively to Conscientiousness (for each, $r = -.16$, $p < .05$), a finding consistent with previous research (Ackerman & Heggestad, 1997; Moutafi, Furnham, & Crump, 2003).

Discussion

Each sample in Study 2 used different methods of assessing the convergent validity of the APM-18. Sample 1 focused on relationships between the APM-18 and other standard measures of intelligence and academic achievement (e.g., verbal intelligence tests, self-reported GPA and SAT scores); while sample 2 examined the relationship between the APM-18, creativity, self-reported ACT scores, and Big Five personality traits. Both studies confirmed that the APM-18 is related to these measures in a predictable manner. Generally speaking, both the APM-18 and the embedded APM-12

showed the same pattern of correlations with the other measures used in these studies. However, the higher internal consistency of the APM-18 suggests that it may be better at detecting individual variation in *g*.

Conclusion

Each of the 18 items used in this new APM-18 test were chosen to maintain the progressive difficulty of both the long form (APM-36) and the short form (APM-12). Unsurprisingly, while the APM-18's reliability was lower than that of the APM-36, it was higher than that of the APM-12 developed by Arthur and Day (1994). Further, the patterns of correlation with other measures of intelligence are virtually identical to the APM-12, which has, in previous studies, been shown to mimic the APM-36 results (Arthur & Day; Arthur et al. 1999). Combined with an average administration time of 17.53 minutes (25 minutes maximum), these findings suggest that the APM-18 may work well as a compromise for researchers who want a quite accurate measure of general intelligence in a quite short amount of time. The cross-validation in the three samples reported here is an initial attempt to collect normative data for the APM-18. Our results may generalize only to other college students. However, the APM-18's short administration time, high internal consistency, reasonable validity, and ease of administration by paper and pencil in large college classroom settings, make it ideal for behavioral science studies where researchers want a reasonably fast, accurate intelligence score as part of a larger questionnaire battery.

References

- Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology*, 77, 598-614.

- Ackerman, P. L., & Heggstad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, 121(2), 219-245.
- Arthur, W., Jr., & Day, D. V. (1994). Development of a short form for the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement*, 54, 394-403.
- Arthur, W., Jr., Tubre, T. C., Paul, D. S., & Sanchez-Ku, M. L. (1999). College-sample psychometric and normative data on a short form of the Raven Advanced Progressive Matrices Test. *Journal of Psychoeducational Assessment*, 17, 354-361.
- Barton, K., Dielman, T. E. & Cattell, R. B. (1971). The prediction of school grades from personality and IQ measures. *Personality*, 2(4), 325-333.
- Brown, M. W. (1994). Cognitive, interest, and personality variables predicting first-semester GPA. *Psychological Reports*, 74(2), 605-606.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1-22.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Dyer, E. D. (1987). Can university success and first-year job performance be predicted from academic achievement, vocational interest, personality and biographical measures? *Psychological Reports*, 61(2), 655-671.
- Forbes, A. R. (1964). An item analysis of the Advanced Matrices. *British Journal of Educational Psychology*, 34, 223-236.
- Hamel, R., & Schmittmann, V. D. (2006). The 20-minute version as a predictor of the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement*, 66, 1039-1046.
- Jackson, D. N., & Rushton, J. P. (2006). Males have greater g: Sex differences in general mental ability from 100,000 17- to 18-year-olds on the Scholastic Assessment Test. *Intelligence*, 34, 479-486.
- McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52(5), 509-516.

- Miller, G. F., & Tal, I. R. (in press). Schizotypy versus openness and intelligence as predictors of creativity. *Schizophrenia Research*.
- Moutafi, J., Furnham, A., & Crump, J. (2003). Demographic and personality predictors of intelligence: A study using the NEO personality inventory and the Myers-Briggs type indicator. *European J. of Personality*, 17(1), 79-94.
- Mouw, J. T. & Khanna, R. K. (1993). Prediction of academic success: A review of the literature and some recommendations. *College Student Journal*, 27(3), 328-336.
- Owen, K. (1992). The suitability of Raven Standard Progressive Matrices for various groups in South Africa. *Personality and Individual Differences*, 13(2), 149-159.
- Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology*, 41(1), 1-48.
- Raven, J. C. (1941). Standardisation of progressive matrices. *British Journal of Medical Psychology*, 19, 137-150.
- Raven, J. C. (1989). The Raven Progressive Matrices: A review of national norming studies in ethnic and socioeconomic variation within the United States. *Journal of Educational Measurement*, 26(1), 1-16.
- Raven, J., Raven, J. C., & Court, J. H. (1993). *Raven manual section 1: General overview*. Oxford: Oxford Psychologists Press.
- Raven, J, Raven, J.C. & Court, J. H. (1997). *Mill Hill Vocabulary Scale: 1998 Edition*. Oxford: Oxford Psychologists Press.
- Rushton, J. P., Cvorovic, J., & Bons, T. A. (2007). General mental ability in South Asians: Data from three Roma (Gypsy) communities in Serbia. *Intelligence*, 35(1), 1-12.
- SAS Institute, Inc. (1999). *SAS/STAT User's Guide, Version 8, Volumes 1, 2, and 3*. Cary, NC: SAS Institute.
- Spearman, C. (1927). *The abilities of man*. London: Macmillan.

Tables

Table 1
Item-Total Correlations and Item Difficulty for the APM-18

	Item-total correlation	% correct	Scale α if deleted
1	.3228	.9289	.7798
2	.3871	.9021	.7754
3	.4160	.8483	.7734
4	.4957	.8168	.7678
5	.3896	.8310	.7752
6	.3973	.7899	.7747
7	.3599	.5861	.7772
8	.3975	.6019	.7747
9	.4800	.5071	.7689
10	.4273	.4801	.7726
11	.2512	.3612	.7846
12	.3214	.3207	.7799
13	.4188	.4739	.7732
14	.3480	.3175	.7781
15	.3211	.2401	.7799
16	.3357	.3144	.7789
17	.2123	.2101	.7872
18	.2734	.1974	.7831

Table 2

Descriptive Statistics for Measures of Intelligence in Sample 1 (N = 193)

	Total		Male		Female	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
APM-18	10.68	3.25	11.07	3.13	10.31	3.34
Shipley	43.47	6.80	42.96	7.69	43.96	5.82
Shipley Verbal	27.65	5.77	27.29	6.65	27.99	4.80
Shipley Abstraction	31.65	5.04	31.34	5.47	31.94	4.59
Mill-Hill	46.87	6.91	46.51	7.98	47.24	5.68
GPA	3.10	0.59	3.03	0.62	3.17	0.56
SAT	1140.98*	163.66	1179.32	143.49	1099.90	174.66

* indicates mean differences as a function of sex

Table 3

Correlations among APM-18, APM-12, and Other Intelligence and Academic Achievement Measures in Sample 1

	Mill-Hill (n = 183)	Shipley Abst (n = 193)	Shipley Vocab (n = 193)	GPA (n = 175)	SAT (n = 144)
APM-18	.22*	.49*	.12	.17*	.34*
APM-12 (embedded)	.26*	.47*	.16*	.12	.32*

* $p < .05$

Table 4
Descriptive Statistics for Measures of Intelligence, Academic Achievement, and Personality in Sample 2 (N = 229)

	Total		Male		Female	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
APM-18	9.53*	3.57	10.29	3.98	9.23	3.36
Verbal Creativity	2.57	.46	2.56	.60	2.57	.40
Drawing Creativity	2.62	.68	2.60	.76	2.63	.65
ACT (n = 129)	23.16	4.20	23.34	4.21	23.08	4.22
Openness	13.83	10.69	13.94	9.84	13.78	11.03
Conscientiousness	11.44	10.78	10.17	11.00	11.93	10.68
Extraversion	12.68	10.69	12.11	11.22	12.91	10.49
Agreeableness	4.74	11.08	3.77	11.27	5.13	11.01
Neuroticism	2.49	12.02	-1.41	12.20	4.04	11.62

* indicates mean differences as a function of sex

Table 5

Correlations among APM-18, APM-12, and Other Intelligence, Academic Achievement, and Personality Measures in Sample 2

	Verbal Creativity (n = 225)	Drawing Creativity (n = 225)	ACT (n = 129)	O (n = 224)	C (n = 226)	E (n = 221)	A (n = 223)	N (n = 225)
APM-18	.35*	.29*	.44*	.26*	-.16*	-.03	-.11	.02
APM-12 (embedded)	.31*	.24*	.45*	.25*	-.16*	-.03	-.07	.00

* $p < .05$