Methodologically Sound: Evaluating the Psychometric Approach to the Assessment of Human Life History [Reply to Copping, Campbell, and Muncer, 2014]

Aurelio José Figueredo, Department of Psychology, School of Mind, Brain and Behavior, College of Science, University of Arizona, Tucson, AZ, USA. Email: ajf@u.arizona.edu (Corresponding author).

Tomás Cabeza de Baca, Health Psychology, Department of Psychiatry, University of California, San Francisco, CA, USA.

Candace Jasmine Black, Department of Psychology, School of Mind, Brain and Behavior, College of Science, University of Arizona, Tucson, AZ, USA.

Rafael Antonio García, Department of Psychology, School of Mind, Brain and Behavior, College of Science, University of Arizona, Tucson, Arizona, USA.

Heitor Barcellos Ferreira Fernandes, Departamentos de Psicologia e de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil.

Pedro Sofio Abril Wolf, Department of Psychology and Centre for Social Science Research, University of Cape Town, Cape Town, Western Cape, South Africa.

Michael Anthony Woodley of Menie, Center Leo Apostel for Interdisciplinary Studies, Vrije Universiteit Brussel, Brussels, Belgium.

Abstract: Copping, Campbell, and Muncer (2014) have recently published an article critical of the psychometric approach to the assessment of life history (LH) strategy. Their purported goal was testing for the convergent validation and examining the psychometric structure of the High-K Strategy Scale (HKSS). As much of the literature on the psychometrics of human LH during the past decade or so has emanated from our research laboratory and those of close collaborators, we have prepared this detailed response. Our response is organized into four main sections: (1) A review of psychometric methods for the assessment of human LH strategy, expounding upon the essence of our approach; (2) our theoretical/conceptual concerns regarding the critique, addressing the broader issues raised by the critique regarding the latent and hierarchical structure of LH strategy; (3) our statistical/methodological concerns regarding the critique, examining the validity and persuasiveness of the empirical case made specifically against the HKSS; and (4) our recommendations for future research that we think might be helpful in closing the gap between the psychometric and biometric approaches to measurement in this area. Clearly stating our theoretical positions, describing our existing body of work, and acknowledging
their limitations should assist future researchers in planning and implementing more informed and prudent empirical research that will synthesize the psychometric approach to the assessment of LH strategy with complementary methods.

**Keywords:** Life History Strategy, psychometrics, biometrics, demography, methodology, Brunswik-symmetry, evolutionary psychology.

---

**Introduction**

A recently published article by Copping, Campbell, and Muncer (2014), hereinafter referred to as “the critique,” has attempted to critically examine the psychometric approach to the assessment of LH strategy that is being increasingly utilized by many evolutionary researchers (e.g., Figueredo, Cabeza de Baca, and Woodley, 2013). The critique presents a number of conceptual and statistical concerns regarding this psychometric approach, broadly summarized as follows:

[Life history theory] originally examined objective, biological events across species (e.g., growth rate, offspring number, body sizes; see Pianka, 1970 for examples). Inventories such as the Mini-K, ALHB, and HKSS focus on factors consistent with a hypothesized “K-oriented” lifestyle (e.g., religiosity, wellbeing, social support, and community engagement). It is difficult to establish how well these personality and lifestyle variables independently measure an individual’s strategy without first validating them against objective life history events. Measures of current wellbeing, integration into the community, and perceived neighborhood safely tell us little about how they would contingently translate into fitness returns. If the HKSS, ALHB, or Mini-K are adequate reflections of an individual’s strategy, high scores should correlate with critical developmental events such as a later pubertal onset, delayed sexual onset, and fewer lifetime sexual partners. Yet these crucial variables are rarely tested in relation to psychometric life history indicators. (p. 204)

We consider the main question that they raise to be worthy of scientific inquiry. Questionnaires that are purported to measure specific constructs should be subjected to validation processes that review the integrity of a measure through multivariate methods, as consistent with theory (e.g., construct validation; see Campbell, 1959). That being said, the psychometric approach has done so by exploring the surrounding nomological net in which they are embedded (Cronbach and Meehl, 1955) and documenting their position within an intricate web of related constructs constructed over years of research (see Figueredo, Cabeza de Baca, et al., 2013). Many of the concerns presented by the critique have been or are currently being addressed within our line of research (e.g., developmental questions, sex differences, cross-cultural validity, nonverbal behavior).

This response to the critique is broadly organized into four main sections: (1) a brief review of psychometric methods for the assessment of human LH strategy, in which we expound upon the essence of our approach; (2) our theoretical concerns regarding the critique and associated conceptual clarifications, in which we address the broader issues raised by the critique for the proper understanding of the hierarchical nature and latent structure LH strategy; (3) our statistical concerns regarding the critique and associated
methodological clarifications, in which we examine the validity and the persuasiveness of the empirical case made specifically against the HKSS; and (4) our recommendations for future research that we think might be helpful in closing the gap between the psychometric and biometric approaches to measurement in this area. We hope that clearly stating our theoretical positions, describing our existing body of work, and acknowledging the limitations of both should assist future researchers to plan and implement more informed and prudent empirical research with the aim of synthesizing the psychometric approach to the assessment of LH strategy with other complementary methods.

Whereas the purported goal of the critique was to test for convergent validation and examine the psychometric structure of the HKSS (Giosan, 2006), the conclusions of the critique with respect to the HKSS were conflated with the purported state of the entire psychometric approach to the assessment of human LH strategy. The conclusions drawn in the critique are generalized beyond the particular measure directly examined. We have therefore decided to address their concerns by using either our own data or parallel results from research that applied better-validated measures: the Arizona Life History Battery (ALHB) and the Mini-K Short Form.

It is also important to note that we are not in any way opposed to the proper use of the conventional biometric indicators of LH championed by the critique. Instead, we completely agree with the critique that, ultimately, all of these measures (regardless of methods) must somehow predict actual LH traits, such as growth, mating, reproduction, and longevity, even though we discuss the various methodological precautions that are needed in so doing. We are also dubious of certain practices that appear to pervade the typical implementation of the conventional biometric approach, such as placing excessive reliance on single-item verbal self-reports to provide data for those critical life events, as will be further detailed below. For example, our research group has attempted to construct multi-item scales out of such single-item self-reports and found that they lack internal consistency, probably due to low item reliabilities as well as validities. Nevertheless, we encourage future LH researchers to pursue the identification of relationships among biometric and psychometric indicators of life history, while cautioning that embarking on such an endeavor requires prudence in its execution, and the limitations we describe must be taken into consideration before implementing or at least interpreting an empirical study. In the fourth main section of the present response, we offer constructive suggestions for successfully testing hypotheses linking the biometric and the psychometric approaches to measurement in LH research.

The Psychometric Approach to the Assessment of LH Strategy

We start by clarifying the use of the word “psychometrics” in specific contrast to “biometrics.” There is no standard definition of psychometrics since its introduction by Sir Francis Galton in 1869. Nevertheless, it is usually taken to mean the theory and methods of measurement applied to psychological constructs that are not directly observable, such as cognitive abilities, behavioral dispositions, attitudes, values, skills, and knowledge. These are usually assessed by means of objective tests, observer ratings, subjective judgments, structured interviews, questionnaires, and other written or verbal responses. Biometrics, in contrast, typically refers to the use of measures related to physical, chemical, or biological characteristics or traits that are, at least in principle, directly measurable, such as the
anatomical or physiological characteristics of individuals (sometimes also called “anthropometrics” in relation to human morphological traits).

We therefore begin our main line of reasoning by briefly reviewing the psychometric approach to the assessment of LH strategy. In both human and nonhuman animals, LH strategies have been assessed by the observation of developmental and physical traits such as gestational length, birth spacing, age of pubertal timing, number of offspring, etc. The critique claims that biometric approaches that include demographic life event data (e.g., anthropometric data, age of sexual maturity, sexual debut, and offspring number) in their LH models are preferable to psychometric approaches that incorporate cognitive and behavioral indicators because they utilize data that are purportedly more “objective.”

As will be detailed below, to posit that “objective life events” are the most valid measures of LH strategy is questionable and ignores the fundamental assumption of LH theory: Resource allocation decision-making is contingent on the ecological challenges the organism is facing. We suggest that instead of creating a false dichotomy between the biometric and psychometric approaches, LH researchers embrace a position where both measures are incorporated into a more inclusive set of measurement and structural models. We take the position that these two kinds of measures assess successive stages within the same causal processes: what Brunswik (1952, cited in Petrinovich, 1979), called the “means” and the “ends” of behavior. The distal achievements (“ends”) produced by functional processes (“means”) can be used when it is theoretically appropriate to do so. Any moderating processes, such as local resource abundance, needs to be specified in this multilevel cascade model to explain when they are not. For example, when studying the Aché, presumably in their ancestral habitat, demographic outcome variables may be entirely valid as indicators of LH strategy, but when studying people living in the context of a modern industrial or post-industrial society (such as contemporary Britain), which represents an environment that introduces novel fitness-relevant adaptive problems (such as the welfare state, legally enforced monogamy, legally enforced child support, contraception, etc.), these demographic outcome measures may be non-trivially complex and equivocal in their interpretation as presumed manifest indicators of heritable and evolved LH strategies.

Thus, the theoretical stance supporting the psychometric approach may be summarized as the need to measure the process of resource allocation within LH strategies and not just the outcomes it produces. This in no way militates against measuring reproductive fitness outcomes as the critique suggests that we should do. However, focusing preferentially on biometric indicators, as implicitly recommended by the critique, would only highlight these fitness outcomes (distal achievements) and ignore the mediating processes (e.g., functional means) that are instrumental in the allocation of material and bioenergetic resources into different components of fitness (see Petrinovich, 1979). Our position is consistent with mainstream evolutionary psychology by considering individual organisms as adaptation executers instead of fitness maximizers (see Alexander, 1990b; Tooby and Cosmides, 1990). Integrating processes (the bioenergetic adaptations being actually executed) into our conceptualization can give us a better characterization of an individual’s LH strategy as the strategy itself is how the organism arrives at the fitness outcomes.
The parameters addressed by conventional biometric approach, such as growth, aging, mating behavior, and parental investment, are thus intermediate ends and not means. The very definition of somatic effort is that growth requires the expenditure of bioenergetic and material resources to achieve. Similarly, successful mating is another achievement or fitness-related outcome which requires the expenditure of mating effort to accomplish. Parental investment is defined as a tradeoff between present and future reproduction (Trivers, 1972), but these fitness gains and losses require the expenditure of parental effort to achieve. Somatic effort, mating effort, and parental effort are essentially behavioral processes, not outcomes like age of puberty, age of first sexual intercourse, number of sexual partners, age at first birth, and number of children produced at any point in time. Psychometric assessments are designed to measure precisely such intervening behavioral processes, the psychological adaptations, and that is precisely what they do best.

The biometric approach to LH assessment is therefore in no way incongruent with the psychometric approach, but rather complementary. Both approaches utilize the general principles of modern LH theory and seek to disentangle the bioenergetic and material allocation tradeoffs organisms make (e.g., somatic effort versus reproductive effort) that produce strategic constellations of physiological, behavioral, and psychological traits to maximize fitness returns within a given environment (see Belsky, Steinberg, and Draper’s [1991] figure on reproductive strategies). For example, this means that organisms should exhibit strategies characterized by distinct and context-dependent patterns arising from allocation trade-offs between somatic effort, which includes growth and physiological maintenance, and reproductive effort, which includes mating and parental effort.

These distinctions are important because environmental mismatch may arise with modern ecologies that produce fitness outcomes contrary to what was ancestrally available in the Adaptively Relevant Environment (ARE; see Irons, 1990). For example, individuals pursuing fast LH strategies historically produced larger numbers of offspring and allocated lower levels of parental effort toward raising their offspring compared to individuals pursuing slower LH strategies. The outcome of these resource allocations during the early modern era (up until the beginning of the 19th Century) was that faster LH strategists ended up with a lower number of surviving offspring as compared with slower LH strategists (see Clark, 2008). Also, whereas it is true that ecologically available resources may limit the fertility of slow LH strategists, they may nonetheless produce a larger number of offspring when resources are temporarily abundant (such as during the settling of the Americas by European colonists), only to revert to lower reproductive rates when this population expansion ceased (see Crosby, 2004). In other words, taking demographic outcome measures at face value and inferring from them just what they mean for LH strategy is not something straightforward. This necessarily implies that the biometric measures purported to be objective measures of LH strategy may not be as unambiguous as they might seem. As Figueredo, Wolf, et al. (2014) noted:

“[M]any researchers deem it safer to settle upon studying the execution of the predicted adaptations (parental effort, nepotistic effort, reciprocity, mutualism, etc.) rather than merely monitoring fitness consequences that might or might not ensue, contingently upon environmental conditions” (p. 151-152).
For example, a modern young male can be genetically predisposed towards a faster life history. However, if he happens to be of lower social status or physical attractiveness, he may not obtain as many mating opportunities as he desires. Furthermore, even if he manages to obtain some, in the age of contraception, this might not directly translate into the production of an increased number of offspring. The widespread use of contraceptive methods also facilitates sexual activity by young, slow LH strategists who might not yet desire offspring. Similarly, the prevalence of legally-enforced monogamy for approximately the past millennium in Western cultures (Wemple, 1981) puts a practical limit (albeit not a complete suppression) on the amount of polygyny an older male may practice, regardless of any genetic predispositions towards faster LH.

What this reasoning logically implies is that the bivariate correlations being reported and proposed by the critique as a model for future validations of assessments of human LH are insufficient to make substantive predictions regarding the relations between fitness outcomes (such as sexual debut and number of sexual partners obtained) and the mediating processes governing LH allocations (such as those behavioral dispositions better assessed by psychometric methods). In the case of hypothetical young male scoring low on the HKSS, one also needs to test for interactions of his HKSS score with independent measures of Physical Attractiveness, Resource Holding Potential, and overall Mate Value (which also include personality dimensions). Furthermore, in cross-cultural research, one must also consider the various social and cultural restrictions under which young people often develop that function to prevent or interfere with early sexual activity, especially if that individual were female.

Thus, we should not necessarily expect that the main effects of LH indicators like the Mini-K or the HKSS will be statistically significant or very large in magnitude when predicting such fitness outcomes (e.g., critical LH events such as sexual debut), unless various other conditions permitting the direct expressions of these heritable dispositions are also met. This is simply because one cannot even have a “sexual debut” without a sexual partner, and that condition is often not satisfied for all individuals out in the real world.

Similarly, a person’s genetic predisposition, or epigenetic bias, towards developing a certain LH strategy is not the only source of variability in pubertal timing. Environmental components such as malnutrition, disease, father absence/stepfather presence, history of mood disorder in the mother (Ellis and Garber, 2000), altitude, and exercise also play a role. Ellis (2004) makes a compelling argument that “from a life history perspective, there is no single answer to the question of when puberty should occur” (p. 948). Ellis also states that there are fitness consequences to variation in pubertal timing that are environmentally-contingent, and that, because of this, natural selection would not favor genetically canalized predispositions but should instead favor the ability to calibrate reproductive strategies in response to prevailing environmental conditions. As pubertal timing calibrates to environmental conditions, linking LH strategies with puberty is a very complex problem. Some of us, for example, are currently analyzing data from an ongoing longitudinal study spanning 10 years using a representative sample of participants who grew up in the Cape Town Metropolitan Area, which includes participants from some of the richest neighborhoods in the world to some of the most dangerous and poor neighborhoods. In the aggregate, the patterns make no sense. Some girls are going into puberty early, but don’t reproduce until they are in their late 20s; others go into puberty late but reproduce soon after. Not until you take the local conditions into account does any of that make sense.
Consider two hypothetical participants as an example: a girl who grows up in a rich suburb and one who grows up in a poor township. Why is it that the girl from the suburb tends to go through puberty earlier than the girl from the township? We propose that it is probably attributable to adequate nutrition. Although she achieves puberty earlier than the girl from the poor township, the rich suburban girl doesn’t reproduce until much later in life, probably due to adequate access to reproductive healthcare. In contrast, the girl from the township achieves puberty later (presumably due to inadequate nutrition) but reproduces earlier (presumably due to inadequate reproductive healthcare and access to contraceptives). Yet another girl from slightly less impoverished area on the proverbial “other side of the train tracks” may go into puberty extremely early. Nutrition is still at play here but the girl probably was provided with just enough calories to not delay pubertal timing. Nevertheless, all the other fitness-relevant indicators might be cueing this girl to reproduce now and to do so prolifically.

We are not saying that using pubertal timing as one indicator of LH strategy among others is not valid. We are saying that relying on any one fallible indicator, however important it might be as a developmental landmark, is ill-advised, hence the need that we discern for data aggregation and latent variable modeling, in spite of its admitted limitations. A study such as that presented by Copping and colleagues (2014), whose aim was assessing the validity of a psychometric measure against such unitary developmental markers, but was performed without controlling for the effects of other known causal influences on those intermediate fitness outcomes (e.g., nutrition, exercise, etc.) does not build a compelling argument that the psychometric measure in question lacks validity.

Furthermore, biometric and psychosocial traits are in a perpetual state of mutual transaction with each other (Bouchard, 1997; Johnson, Turkheimer, Gottesman, and Bouchard, 2010). In certain cases, psychometric antecedents have been used to predict biometric outcomes. For example, a large body of literature has found psychosocial and environmental antecedents of early pubertal timing in girls. Ellis and Garber (2000) reported that the association between maternal psychiatric history and early pubertal timing in girls was mediated by higher levels of maternally-reported romantic conflict and father absence within the household. In other research, harsh and coercive parenting from the father and lower levels of parental cooperation—psychosocial variables—predicted earlier pubertal timing (Ellis and Essex, 2007; Ellis, McFadyen-Ketchum, Dodge, Pettit, and Bates, 1999). Further, father absence, one of the psychosocial and environmental antecedents tested in pubertal timing, also predicted early sexual debut (Ellis et al., 2003). In other cases, researchers have found the reverse, with biometrically-assessed antecedents predicting psychosocial outcomes. For example, Belsky’s (1997) model of differential susceptibility posits that an individual’s genome may be a buffer or an “amplifier” to environmental and social input. Individuals with differential susceptibility genes that reside in a stress-producing environment (a biometric antecedent) disproportionally manifest high-risk behaviors and emotional problems (for a brief review of the effects of differential susceptibility, see Belsky, Bakermans-Kranenburg, and van IJzendoorn, 2007).

The results from these two reviewed bodies of literature are consistent with the prediction by LH theory that faster LH strategists should invest more energy in reproductive effort, and particularly in mating effort, while minimizing investments in growth and maintenance. This implies that the pattern among slow LH individuals will emerge as one of higher somatic effort and lower mating effort. These results also show
quite unequivocally a mutually influential (transactional) relationship among psychosocial and biological traits within an individual and with the environment. It would be a disservice toward the understanding of human LH strategies to ignore or downplay the importance of psychosocial and cognitive traits measured by psychometric assessments of LH strategy. Humans are distinct from non-human animals in that we can provide self-report measures of biological, demographic, cognitive, behavioral and personality traits to capture a more nuanced and well-measured factor of life history, as users of the psychometric approach seek to do.

Finally, this critique of the psychometric approach expresses apprehension regarding the usage of personality measures as indicators of LH strategies. As with other fitness components, however, the development and behavioral enactment of personality traits require the expenditure of material and bioenergetic resources for the observable indicators of personality to outwardly manifest (Figueroedo, Woodley of Menie, and Jacobs, in press). For example, Agreeableness is a personality trait that requires an outward behavioral manifestation consistent with the verbal adjectives participants utilize in both self- and peer-report measures. To be considered agreeable, by either oneself or by others, one must allocate time and energy towards prosocial and altruistic behaviors when dealing with conspecifics, as well as towards developing the interpersonal and emotional self-efficacies needed to identify the needs and desires of others (for the optimal deployment of these prosocial and altruistic investments). As a result, being agreeable and seeking to ingratiating oneself among groups of conspecifics is not without certain bioenergetic and material cost. Actively self-selecting into and occasionally even constructing situations and environments that provide the settings for one’s agreeable behavior will also necessarily reduce the temporal, material, and bioenergetic resources that are allocated towards other fitness-enhancing domains. The same general argument can be made for any of the other Big Five personality traits (and has already been in Figueredo et al., in press). If the “possession” of personality traits consumes resources to produce elaborate behavioral extended phenotypes, we can only conclude that personality traits are components of a higher-order LH strategy ancestrally selected to optimize the tradeoffs between the costs and benefits of the behaviors. As a result, any organismic resources allocated toward personality traits should be invested in a manner that ancestrally produced higher fitness returns on investment by strategically distributing resources across multiple fitness domains. The magnitudes of these fitness returns, in relation to the costs of the investments, may be highly contingent on the ecology in which the individual resides (for a more explicit model of these quantitative relations, see Figueredo et al., in press).

The critique insists that the constructs sampled in psychometric measures of LH strategy cannot stand as proxies for “true” LH traits, such as pubertal timing, age at sexual debut, and number of sexual partners (which are the “classic” LH traits). Nevertheless, Gallup (2009) found that a composite well-being score correlated with life expectancy ($r = 0.71$), infant mortality rate ($r = -0.58$), and teenage pregnancies ($r = -0.38$), traits that the critique evidently considers “factors consistent with a hypothesized ‘K-oriented’ lifestyle (e.g., religiosity, wellbeing, social support, and community engagement)” (p. 204). Similarly, moderate to strong associations are usually found in nonhuman animals when examining the relationship among traditional LH traits and putative cognitive-behavioral indicators of life history, such as experimental tests of willingness to wait for delayed rewards (Stevens, 2014) as well as open field experimental assessments of boldness in

Reply to Copping, Campbell, and Muncer (2014)
response to risky situations (for a review, see Biro and Stamps, 2008). Furthermore, as pointed out by Stearns (1992) and others, there are many more LH traits than are usually studied in biology. In this view, body size, growth pattern, age at maturity, number of offspring, and longevity are the main traits studied, and there are three or four tradeoffs (such as mating versus parenting effort) that are usually the focus of LH studies, whereas “at least 45 trade-offs are readily defined between life history traits” (p. 72). That does not mean that certain LH traits serve as mere “proxies” for others, but that there are many more LH traits than are dreamt of in some philosophies, and that these traits are correlated tactical elements within the coordinated strategic systems that we refer to as life history.

The inclusion of psychological and cognitive variables in the patterns of human LH evolution is not unique to scholars who rely primarily upon the psychometric approach to the assessment of LH traits. For example, Alexander and his colleagues (Alexander, 1990a; Flinn, 2006; Flinn, Geary, and Ward, 2005) have proposed the Ecological Dominance/Social Competition model, whereby human evolution and brain development were primarily driven by intraspecific competition (including between social groups) in population-dense ecologies. They posited that after humans were able to master what Darwin (1859) called the “hostile forces of nature” and deter more physically powerful predators (such as leopards), the next ecological challenge that arose was interacting (both competing and cooperating) with conspecifics. As a result of these social selective pressures, the mastering of context-contingent social skills became a valuable tool for survival and reproduction. The social demands of our environment led to systematic and coordinated alterations to the LH traits of humans (e.g., reduced canine size, reduced sexual dimorphism, prolonged infant development) and their associated neuroanatomy, having implications for human cognitive abilities and propensities, including foresight and planning (see Flinn et al., 2005, Table 1). Thus, we believe that whilst the inclusion of cognitive variables may not have been necessarily universally endorsed among LH researchers working with nonhuman participants, excluding them from inquiry would limit the evolutionary scenario many researchers before us have sought to reconstruct.

The particular traits favored by the critique are the focus of many LH studies in biology mainly because they are easier to measure than other LH traits, and some of them have also been included in psychometric measures such as the Mini-K. That does not mean that those traits usually studied in biology are necessarily “better” LH indicators than, say, insight, planning, and control. Although self-reports in humans might not be perfectly reliable or valid, at least they are easily available and thus amenable to psychometric study. Due to a lack of verbal self-reports, such traits are just very hard to assess in flies, deer, and lizards.

**Theoretical Concerns and Conceptual Clarifications**

What follows are our responses to various theoretical problems posed by the critique and the associated clarifications needed to resolve the conceptual difficulties involved.

According to the critique, advocates of the psychometric approach are removed from the contemporary advances occurring within LH. The critique asserts that the early paradigms of the r-K continuum, rooted in density-dependent models and emphasizing the carrying-capacity of the environment, are central to the theories of present
psychometrically-oriented human LH researchers. The critique instead emphasizes the relevance of local age-specific mortality rates in relation to population density in ecological conditions. Although evolutionary psychologists have utilized LH theory in their conceptualization of individual differences in many traits (e.g., Kaplan and Gangestad, 2005), there exist several different stages in the historical development and elaboration of LH theory, and most contemporary human LH theorists are by no means archaic in their empirical approach toward human reproductive strategies. Instead, human LH theory has moved to more complex paradigms, incorporating social development (e.g., Belsky et al., 1991; Ellis, 2004; Hrdy, 1999), context and age-dependency (e.g., Ellis, Figueredo, Brumbach, and Schlomer, 2009), behavioral endocrinology and epigenetics (e.g., Del Giudice, Ellis, and Shirtcliff, 2011; Flinn, 2006; Meaney, 2010), stress responsivity (e.g., Del Giudice, Hinnant, Ellis, and El-Sheikh, 2012), and sex differences (e.g., Del Giudice, 2009). Thus, we acknowledge that human LH strategy is highly variable and greatly dependent on a multitude of contextual factors. Although we also acknowledge that LH strategies have a heritable component (Figueredo, Vásquez, et al., 2004; Figueredo and Rushton, 2009), asserting that we believe only “the existence of heritable clusters” (Copping et al., 2014, p. 201), as is done in the critique, is incorrect and disregards an even larger proportion of our work that is specifically related to environmental influences.

For example, one earlier conceptualization of LH theory, called Differential-K Theory, was originally proposed by Rushton (1985), who posited that morphological, psychological, and behavioral trait differences existed both at the inter-individual and at the inter-ethnic levels. Whereas Differential-K Theory was based upon an early conceptualization of LH theory more prevalent at the time, the field of LH evolution has progressed substantially since then. For example, more recent findings regarding systematic inter-ethnic differences in LH strategy have shown that individual differences within groups explain over 90% of the variance in LH traits, whereas self-reported ethnic categorizations explain less than 10% of the variance in the individual domain-specific indicators as well as in the multivariate K-Factor aggregate (e.g., Figueredo, Vásquez, Brumbback, and Schneider, 2004, 2007).

Although we have retained the name K-Factor as a label for this latent construct, density-dependent selection for slower LH is not central to our model. In other words, the psychometric K-Factor does not directly correspond to the products of K-selection as described in Pianka’s (1970) r-K theory, but is instead multicausal in evolutionary origin, as per contemporary LH theory (e.g., Charnov, 1993; Roff, 2002; Stearns, 1992; but for a defense of the continuing relevance of density-dependent selection as an important factor in LH evolution, also see Reznick, Bryant, and Bashey, 2002). This distinction has been explained many times in our publications (e.g., Ellis et al., 2009), and criticisms of the psychometric approach based upon our use of K as a convenient symbol present all the appearance of attacking a “straw man,” simply because we completely agree with the critique that LH strategies are multicausal. As we have stated explicitly (Figueredo, Cabeza de Baca, et al., 2013):
by Rushton (1985) for the specific implications of LH strategy in our species. (p. 253)

Furthermore, the psychometric approach to the assessment of human LH strategies emphasizes individual differences. This approach suggests that individual humans occupy many different loci along a continuum of slower to faster LH strategies, consistent with the species-typical strategy generally employed by humans (Ellis et al., 2009). We should therefore fully expect to see differences both within and between human groups (Ellis et al., 2009), simply because there is no mammalian species extant with a comparable worldwide distribution that has been found to be genetically homogenous across geographically isolated regions.

One pair of groups between which we should expect to see differences is among men and women (Trivers, 1972). The Copping et al. (2014) critique made the important recommendation that sex differences should be tested when examining LH strategies, but asserts that within the literature relying on the psychometric approach to assessment, “sex differences are rarely considered” (p. 203). This is incorrect, as several recent publications have in fact reported systematic sex differences of between one-quarter to one-third of a standard deviation between sexes, with males being predictably “faster” than females (Figueroedo, Andrzejczak, Jones, Smith-Castro, and Montero-Rojas, 2011; Figueredo, Cabeza de Baca, et al., 2013; Gladden, Figueredo, Andrzejak, Jones, and Smith-Castro, 2013; Gladden, Sisco, and Figueredo, 2008). Alternatively, what the authors of the critique might mean by this point is that the same measures were used to assess LH strategy in both males and females, and not that differences in their relative levels were not examined. If that is indeed what they mean, then we respectfully disagree. One cannot perform valid systematic comparisons among sexes (or any other groups) using two different metrics (such as centimeters for females and inches for males, from which one might conclude that the average female is taller). Psychometricians use the same measures across sexes specifically in order to be able to make such comparisons.

Another limitation of the psychometric approach that was asserted in the critique was with regards to cross-cultural validity of the psychometric measures of life history:

Samples used in recent works are almost exclusively from the US, and so cross-cultural validity is lacking. This concern about sampling is also true of other research using psychometric indicators of life history strategy. (p. 203)

While the HKSS has been applied in only 12 studies to date, as noted in the critique, research on the Arizona Life History Battery (and its Short Form, the Mini-K), has produced quite a number of cross-cultural studies in recent years in geographically diverse places such as Chile, Costa Rica, Israel, Mexico, Poland, Singapore, Sweden, the Netherlands, and the United Kingdom (e.g., Abed et al., 2012; Buunk and Hoben, 2013; Buunk, Pollet, Klavina, Figueredo, and Dijkstra, 2009; Cabeza de Baca, Figueredo, and Ellis, 2012; Cabeza de Baca, Sotomayor-Peterson, Smith-Castro, and Figueredo, 2014; Egan et al., 2005; Figueredo et al., 2011; Figueredo, Cabeza de Baca, et al., 2013; Figueredo, Tal, McNeill, and Guillén, 2004; Figueredo and Wolf, 2009; Frías-Armenta et al., 2005; Frías-Armenta, Valdez-Ramirez, Nava-Cruz, Figueredo, and Corral-Verdugo, 2010; Gaxiola-Romero and Frías-Armenta, 2008; Jonason, Li, and Czarna, 2011;
In specific relation to the HKSS, which again has been applied in only 12 studies to date, the critique correctly notes that “the utility of this scale as a measure of life history strategy is difficult to determine without a large and representative sample of the general population” (p. 203). However, while this might indeed be true of the HKSS, the critique continues by generalizing that assertion to the rest of the research within this area: “This concern about sampling is also true of other research using psychometric indicators of life history strategy” (p. 203).

The problem with this generalization is that, again, research on the Arizona Life History Battery (and its Short Form, the Mini-K) and related psychometric measures of human LH strategy has produced quite a number of population-representative studies in recent years (e.g., Brumbach, Figueredo, and Ellis, 2009; Brumbach, Walsh, and Figueredo, 2007; Figueredo and Rushton, 2009; Figueredo, Woodley, Brown and Ross, 2013; Figueredo, Vásquez, et al., 2004, 2007; Woodley, Figueredo, Brown, and Ross, 2013). Even without considering those more nationally-representative samples, the “general population” is more properly considered a metapopulation made of the “special populations” to which the critique takes such exception, implying that a sufficient number and combination of special populations support the claim of greater generality (Figueredo, Black, and Scott, 2013a, 2013b). One recent study (Figueredo, Woodley et al., 2013), for example, cross-validated the results for the Strategic Differentiation-Integration Hypothesis across samples from (1) the National Longitudinal Survey of Youth (NLSY), with participants ranging from late childhood to late adolescence (12–18 years old), (2) two samples of college undergraduates (18–25 years old) with different sexual compositions, and (3) the MIDUS Sample of adults (25–75 years old). We do not believe that this sampling plan missed any natural age group of substantial size.

Furthermore, a meta-analytic construct validation was recently published (Figueredo, Wolf, et al., 2014) examining both the convergent and the nomological validities of the Mini-K (the 20-item Short Form of the Arizona Life History Battery; Figueredo, Vásquez, et al., 2006) by correlating it with other relevant LH constructs. Among the limitations of the study that were examined was whether there were significant restrictions of range in the psychometrically-assessed LH strategies of English-Speaking North American Undergraduate College Student Samples used in those analyses with respect to the general English-Speaking North American population as a whole. Variances for the general English-Speaking North American population on scales that had been incorporated into the ALHB were obtained from the National Survey of Midlife Development in the United States (MIDUS Survey) data on scales for direct comparison with those included in the ALHB and used on the meta-analytic student samples; variances for the general English-Speaking North American population on the Mini-K Short Form were obtained from a recent internet-based study that collected Mini-K samples from all over North America (Fernandes, Kruger, Natividade, Hutz, and Kennair, 2012) for direct comparison with the variances on the student Mini-K samples that had been used in the meta-analytic samples. In both cases, the answers were quite similar: there were no
significant differences in variance among the undergraduate and the general population samples. This means that the distribution of psychometrically-assessed LH strategy (K-Factor) scores of the large number of “convenience samples” used in this meta-analytic construct validation were highly reflective of those of the general population from which they were ultimately derived. Owing to the fact that this meta-analytic construct validation includes a non-trivial proportion of the extant literature using the psychometric approach to the assessment of human LH strategy, we may conclude that the blanket assertion that the samples used are grossly unrepresentative is simply incorrect.

Furthermore, as explained by Cabeza de Baca, Sotomayor-Peterson, Smith-Castro, and Figueredo (2014):

Students cannot be used incautiously as surrogates for general populations. Nevertheless, recent analyses by Flere and Lavrič (2008), based on the World Values Survey data on mean values of four sociologically and psychologically relevant measures comparing between national and student samples of 23 countries, showed that comparisons of student samples are reliable predictors of general cross-cultural differences. This is because we can to some extent infer the cross-cultural differences in the general populations from those observed in student samples, given the known systematic differences between student and nonstudent samples that are fairly generalizable across cultures that have student populations (Henrich, Heine, and Norenzayan, 2010). (pp. 549–550)

Cross-cultural studies conducted by our own research group using psychometric measures, such as the Mating Effort Scale (MES), have also found relatively trivial differences between student and non-student samples of various different nationalities (e.g., Egan et al., 2005).

Given that the critique reports that the HKSS has only been utilized in 12 published studies so far, we are surprised that Copping and his colleagues would choose to select a scale for their critique with such a small body of published research associated with it, whereas a comparable psychometric Short Form like the Mini-K had already been used in the 34 completed studies by 2009 that were included in the recent meta-analysis, to which more studies have been added since. Dunkel and Decker (2010), among others that are cited in the critique, show that the Mini-K tends to better predict expected longevity, short-term mating orientation, future time perspective, and the general factor of personality (GFP) than the HKSS does. The selection by Copping and colleagues of the HKSS, which represents the least validated psychometric scale for LH assessment, as the main target of the critique reflects a quite puzzling and perhaps unintentionally misleading choice.

The critique also claims that the psychometric approach routinely ignores important “contingent relationships,” as the authors refer to them. This is also factually inaccurate, as best illustrated by the work of Brumbach et al. (2009), which used data from the National Longitudinal Study of Adolescent Health (Add Health) to develop a longitudinal structural equations model of developmental risk factors that tracked participant LH strategy through different stages of development. A more recent study (Figueredo, Cabeza de Baca, and Black, 2014) tracked the temporal stability of the psychometrically-assessed LH strategies of participants across both Wave 1 and Wave 2 of the MIDUS Survey, taken 10 years apart.
Paradoxically, the critique calls for more longitudinal studies (a sentiment which we endorse) but offers a cross-sectional one in purported refutation of this corpus of research.

In the same vein, the critique claimed that psychometric measures of human LH strategy generally “contain a blend of items assessing current and past environments, relationships with parents and offspring, personality, and lifestyle” (p. 204). They argue that such a measure would be “problematic” for evolutionary developmental researchers utilizing Psychosocial Acceleration Theory (Belsky et al., 1991) because it conflates “factors which are proposed to canalize developmental strategy” with “mediators of strategy” and with “LH outcomes and correlates” (Copping et al, 2014, p. 204). Measures such as the Arizona Life History Battery do indeed include domain-specific social and behavioral facets taken from an individual’s past and present life. This is in no way problematic from the standpoint of psychosocial acceleration theory. Psychosocial acceleration theory argues that early environments, particularly family environments, produce alterations in the child’s physical development and later social and sexual behavior. Belsky and colleagues propose a developmental process model that begins with family context (e.g., marital relationships and financial resources) that may either influence or be genetically correlated with (by means of active gene-environment correlation) the type of childrearing environment that parents construct for their children. A child’s psychological/behavioral development (e.g., attachment and interpersonal orientation) is presumably shaped by the type of childrearing that the child receives from his or her parents. This means that if the child grew up in a family with harsh parenting, the child would develop a worldview reflective of an upbringing that equips the child with cognitive/behavioral tools in anticipation of their likely future. The next step in the Psychosocial Acceleration model involves alterations to somatic development. In this step, environmental and psychosocial antecedents modulate tempo and timing of puberty. Whether a child reaches puberty earlier or later, this leads to alterations in reproductive strategy. According to the Psychosocial Acceleration model, girls who reached puberty early would exhibit early onset of sexual debut, an unrestricted sociosexual orientation, unstable pair bonds, and decreased parental effort.

The psychosocial acceleration theory posited by Belsky and colleagues (1991) and further elaborated by Ellis et al. (2009) provides a testable environmental pathway for developmentalists to examine. Developmentalists, depending on their area of interest, can test pieces and components of the model (e.g., Belsky, Schlomer, and Ellis, 2012; Simpson, Griskevicius, Kuo, Sung, and Collins, 2012), as the critique describes. The psychosocial acceleration model, however, also provides excellent psychometric indicators of LH strategy. If one is trying to capture an overall assessment of LH strategy in the type of adult participant typically utilized in psychometric studies, then the components present in the psychosocial acceleration pathway provide researchers with an excellent “snapshot” of an individual’s LH development.

We have always wholeheartedly endorsed the existence of a developmental sequence among some or all of the convergent indicators of the K-Factor. This is because we theorize that such continuing influence of a heritable coordinating strategy is required by biologically-based models of the evolution of developmental plasticity (such as those of Figueredo, Hammond, and McKiernan, 2006; West-Eberhard, 2003). The ALHB was, in fact, initially constructed with just such a developmental sequence implicit in its structure. Specifically, it was partially inspired by The Bioecological Model of Human Development.
proposed by Bronfenbrenner and Morris (2006), on which we have based other evolutionary work (e.g., Figueredo Brumbaeh, et al., 2007). Multivariate psychometric assessments so designed can therefore be used simultaneously by: (1) developmentally-minded evolutionists to test posited pathways and disentangle mediators, moderators, and outcomes; and (2) personality and individual differences-oriented researchers to create factors based on the indicators and test alternative measurement models for human life history. What would be optimal, of course, is if such individuals could work together to produce a more integrated and coherent body of multidisciplinary knowledge.

One such integrative set of multivariate analyses was performed to evaluate the proposition that the construction of a common factor model (such as our K-Factor Model based on the ALHB) was somehow inconsistent with that of a developmental sequence (such as that proposed by Adaptive Calibration Theory). Garcia, Cabeza de Baca, Sotomayor-Peterson, Smith-Castro, and Figueredo (in preparation) recently used a large data set on the ALHB to test a plausible Bronfenbrenner-based developmental sequence among all of its subscales, with the exception of the Mini-K. As the Mini-K is a Short Form for the entire K-Factor, and not any domain-specific area of resource allocation (as represented by each of the other scales), we used it as a proxy for the general K-Factor to specify a “hybrid” model, containing both kinds of effect (factor loadings and developmental pathways). Our best meta-analytic estimate of the disattenuated population-level validity of the Mini-K with respect to the K-Factor comes in at $\rho = .91$ ($p<.05$; Figueredo, Wolf, et al., 2014), justifying its use as a convenient proxy for this purpose; the heritability of the Mini-K has recently been established elsewhere (Woodley of Menie and Madison, 2015), as distinct from that of the entire K-Factor, which had been published ten years earlier (Figueredo, Vásquez, et al., 2004; see also Figueredo and Rushton, 2009)

This hybrid model fit the data substantially better ($\chi^2[10] = 23.67*$; $RMSEA = .06$; $CFI = .98$) than alternative (“restricted”) models based on either (1) the common factor alone ($\chi^2[9] = 52.05$, $p<.05$; $RMSEA = .11$; $CFI = .86$) or (2) the developmental sequence alone ($\chi^2[10] = 36.51$, $p<.05$; $RMSEA = .08$; $CFI = .91$). Figure 1 provides a diagrammatic representation of this model. These results support the view that not only is a common source of continuing genetic influence consistent with the proposed developmental sequence, but it is epigenetically necessary for keeping it on track.

Although the psychometric approach composites multiple indicators present in the developmental pathway of the psychosocial acceleration model for the purpose of constructing latent variables, it can also be used to test components of the psychosocial acceleration model. For instance, retrospective self-report research on Mexican, Costa Rican, and U.S. students (Sotomayor-Peterson, Cabeza de Baca, Figueredo, and Smith-Castro, 2013) examined the total amount of parental effort, shared parenting, and positive family emotional climate that the individual experienced as a child. As predicted by psychosocial acceleration theory, slow LH strategists reported having experienced higher levels of total parental effort, more cooperative shared parenting, and higher levels of positive emotional family climate. This means that slow LH strategists recall growing up in environments that were stable, prosocial, and cooperative. Follow-up retrospective research on Mexican and Costa Rican students (Cabeza de Baca et al., 2014) also found that slow LH strategists reported higher levels of positive family emotional climate, consistent with the previous study. Additionally, slow LH strategists reported lower levels of parental harshness. In these studies, independent psychometric measures of behaviors were also
used besides the ALHB, such as the *Parental Effort Scales* (Cabeza de Baca et al., 2012) that directly enumerate the behavioral contributions enacted by the parents towards the child’s welfare.

**Figure 1.** Hybrid Bronfenbrenner-based model of genetically-guided LH development

A concern raised by one reviewer of a previous draft of this paper was that the measures used in the two cited studies to assess slow LH in young adulthood contained items and subscales which referred to the quality of early parent-child relationships, although they did not ask explicitly about the amounts of Parental Effort received or Shared Parenting observed. Nevertheless, we reanalyzed the data from both studies, omitting the potentially confounding subscales, and found that the slow LH factors with and without those specific subscales included were highly correlated (*r* = .92, *N* = 469, *p* < .0001, in the Sotomayor-Peterson et al. [2013] data; and *r* = .96, *N* = 442, *p* < .0001, in the Cabeza de Baca et al. [2013] data). Furthermore, the correlations among the slow LH factors and the other major constructs in these studies were only slightly attenuated by the removal of the affected subscales, remaining statistically significant and in the expected directions.

Nevertheless, the critique makes a valid point in stating that common factor models using global inventories such as the ALHB are relatively agnostic with respect to possible causal relations among its various components. Any common factor model, and not just ours, represents the multiple convergent indicators as effects of a common unobserved influence, which is the latent variable hypothesized. The purpose of such “measurement models” is to measure the phenomenon in question by means of whatever manifest indicators can be systematically associated with it. Nothing in this procedure prohibits researchers from modeling and testing the possible causal relations that might exist among these indicators, other than the spurious relations produced by the common causal influence of the latent factor. Nevertheless, it must be kept in mind at all times that such causal relations, as with all causal hypotheses, are themselves inferential and not directly observable, or known with perfect certainty. In contrast, the authors of the critique apparently presume to know which of the various LH traits being measured are causal to
the others. To a philosopher of science, this degree of confidence would appear remarkable. We do not assert that they are wrong, as this would be as philosophically questionable as asserting that they are right, but merely point out that a common factor model is instead designed to *measure* the various components of a more general LH strategy without taking any strong position regarding which of them might or might not causally influence the others. For example, many of the putative developmental “causes” of certain LH traits, such as those measured in the Family Environment Scale (FES), have been shown to be highly heritable (Rowe, 1994) and might therefore be genetically correlated with the purported outcomes, rather than be a true causal influence upon them. In several studies (e.g., Figueredo, Cabeza de Baca, and Black, 2014; Figueredo, Vásquez, et al., 2004, 2007; Figueredo and Rushton, 2009), our lab has shown that there are high and statistically significant genetic correlations among all the indicators within the ALHB and related inventories, making this more than just a speculative possibility.

Nevertheless, there are many other aspects of this complex psychosocial acceleration theory that are indeed amenable to psychometrically-informed analyses, such as the development of opportunistic interpersonal orientation as a hypothesized result of accelerated LH and reciprocally rewarding interpersonal orientation as a hypothesized result of decelerated LH. Our lab has indeed addressed these issues in various structural equation models, using psychometric assessments such as the *Young Schema Questionnaire* (e.g., see Figueredo, Gladden, and Hohman, 2012) to model the possession of social-cognitive schemata that we have characterized as Antagonistic versus Mutualistic, which roughly reflect the original construct proposed by Belsky et al. (1991).

Evolutionary-development psychologists posit that there are key developmental decision nodes that open and/or close other developmental decisions once the individual reaches the next developmental decision. Over time, a cohesive and unified pattern begins to emerge from the accumulation of developmental decision node outcomes experienced in life (e.g., a crystallized LH strategy). If one is investigating decision nodes in young children or adolescents, there might not be much need for compositing of indicators; however, if a researcher is investigating a developmental question using an older sample, such as college students or older adults, a “snapshot” of the cumulative effects of all prior developmental decisions is both conceptually and methodologically appropriate. Of course, taking multiple “snapshots” (as in longitudinal study) is preferable to taking a single one, as these would better track the effects of each successive developmental milestone. That is because the cumulative effects of developmental decisions taken at these successive nodes are not independent outcomes over an entire lifetime. They are instead like ripples in a pond, with each ripple connected to the next one. One developmental decision does not define an LH strategy.

**Methodological Concerns and Clarifications**

We first take up the various concerns raised in the critique regarding the HKSS itself and then move on to other, better-validated psychometric assessments of human LH. As some of these discussions are rather technical in nature, we have relegated the finer points to an appendix on Statistical Concerns (see Appendix A).

The main theoretical problem with the critique’s examination of the internal structure of HKSS is the implicit expectation that either the HKSS or the Mini-K Short
Form represent a unitary and homogenous factor. Both are instead construed by all as indicators of higher-order factors, which is something that has been reported in numerous publications during the previous decade, sometimes irreverently referred to as the “naughties” (e.g., Figueredo, Vásquez et al., 2004, 2006, 2007). Most recently (Figueredo, Wolf, et al., 2014), this idea was expressed in the following sentence:

…the Mini-K serves as a direct measure of the latent common factor (K) underlying the several convergent scales of the ALHB, and is therefore not a typical “short form” of a single unidimensional scale. (p. 9)

Of course, it is evident that the authors of the present critique do not subscribe to the idea of higher-order LH factors. Nevertheless, it is also important to recall that a hierarchically-organized system of nested common factors representing the latent structure of LH strategy has been found time and time again across class Mammalia and its orders, whether as a whole or in part (Del Giudice, 2014; Dobson and Oli, 2007; Fernandes, Figueredo and Woodley of Menie, under review; Oli, 2004; Promislow and Harvey 1990; Read and Harvey 1989; Rushton, 2004; van Schaik and Isler 2012; but see Appendix A for one seeming anomaly in the literature). We would therefore be very surprised if LH strategy in humans were not also found to possess a latent structure that is similarly hierarchical in organization. As with the psychometric human data, biological LH traits are usually moderately to strongly correlated with social characteristics among mammals and especially among primates.

Members of our research group have recently examined the factor pattern among seven conventional LH traits with a relatively large sample of 120 primate species (Fernandes et al, under review), as compared with the 86 species examined by van Schaik and Isler (2012) and the 41 species examined by Bielby et al. (2007). The specific LH traits sampled were average interbirth interval, gestation length, lactation length, age at female maturity, longevity, and neonatal and adult body mass. Consistent with the results of a principal components analysis computed by van Schaik and Isler (2012), in all our principal axis factor analyses (using prior communality estimation, which discriminates between shared and error variance and is thus unlikely to overestimate shared variance among variables) with an oblique rotation (a more generally appropriate factor analytic method than orthogonal), only one factor was extracted according to Horn’s (1965) Parallel Analysis (see Ledesma and Valero-Mora, 2007, for a discussion on the advantages of this method over alternative methods for the identification of the number of factors extracted).

With raw species-level data, the share of the variance in LH traits explained by the LH factor was 81% (factor loadings ranged from .75 to .95). Using phylogenetic contrasts of the LH traits (after adjusting the branch lengths of the phylogenetic tree for the phylogenetic signals of the LH traits estimated both with Pagel’s λ and Blomberg’s K; see Kamilar and Cooper, 2013 for methodological explanations), 44% to 53% of the variance in the LH traits were explained by the LH factor (factor loadings ranged from .41 to .90). This suggests not only that LH traits currently covary among extant primate species, but that they also coevolved in considerably tight lockstep throughout the primate evolutionary history. We also tested whether the putative allometric effects of adult body mass upon LH traits was responsible for their pattern of coevolution, but the results largely indicated that it is not, as only one LH factor was extracted with the residuals of LH traits, with 30% to
36% of their variance explained by the LH factor after accounting for phylogenetic effects (factor loadings ranged from .38 to .83). These latter estimates of the evolutionary covariance among LH traits are likely to be underestimates however, because as discussed previously (e.g., Roff, 2002; Del Giudice, 2014), residualizing traits against body mass eliminates genuine LH trade-offs; moreover, LH variation is likely to cause variation in body size, and not only the contrary (Harvey and Purvis, 1999; Stearns and Koella, 1986).

A robust fast-slow LH dimension is thus recovered in the mammalian order of which humans are part. Moreover, this unitary common factor was found to positively and significantly predict, at the cross-species level, brain size and various indicators of sociality and social complexity (Fernandes et al., under review). To summarize this point, a common factor of LH traits is frequently estimated, and its association with other variables tested in the biological study of nonhuman animals (e.g., Dobson and Oli, 2007; Gaillard et al., 2005; Johansson, 2000). Common factor models are also used in morphometrics, including variables that causally influence each other allometrically, as the critique proposes that LH traits may causally influence each other. These common factors are then correlated with other variables (e.g., Böhning-Gaese, Schuda, and Helbig, 2003; Schulte-Hostedde, Millar, and Hickling, 2001). Predictive common factors models are also used in microbial biology (e.g., Pennanen et al., 1998), plant biology (Eckert, 2000), and other fields in which biometric, as opposed to psychometric data, are obtained.

In contrast, the critique’s preferred hypothesis that LH traits causally influence each other ontogenetically (as opposed to being spuriously related due to the common causal influence of highly heritable factors) is not shared by most biologists in the field of LH theory, as heritability estimates for LH traits have frequently been found to be quite high and sometimes as high as 1.0 in different species (e.g., Dingle and Hegmann, 1982). Although the critique appears to characterize a latent variable approach as inappropriate where LH traits might causally influence each other, common factor models are widely used in fields (such as morphometrics) where variables are theoretically expected to causally influence each other (in that case, allometrically).

We do not mean to infer from these results that there is one and only one dimension underlying the latent structure of LH strategy, even in primates. As with the correct interpretation of the General Intelligence Factor (Spearman’s g) and the General Factor of Personality (GFP), the pertinent question here is not whether there are multiple lower-order factors or merely one higher-order one, because both levels (and more) necessarily exist in any biological system. As best expounded by Mayr (1982):

> In such a hierarchy the members of a lower level, let us say tissues, are combined into new units (organs) that have unitary functions and emergent properties… At each level there are different problems, different questions to be asked, and different theories to be formulated. Each of these levels has given rise to a separate branch of biology; molecules to molecular biology, cells to cytology, tissues to histology, and so forth, up to biogeography and the study of ecosystems. (p. 65)

This implies that there are always specific variance components that are not shared among the different levels of aggregation in any latent hierarchy. For example, correlations among variables that are mismatched in their relative levels of aggregation within that hierarchy might be misleading with respect to the true magnitudes of the latent associations
among the corresponding constructs at the same levels of aggregation. This phenomenon has been referred to by methodologists as the Principle of Brunswik-Symmetry (Brunswick, 1952).

One of the principal limitations in the reasoning applied by authors of this critique is probably attributable to an evident inattention to the implications of Brunswik-Symmetry for the problem of LH trait measurement and validation. The Brunswik-Symmetry argument emphasizes the need to correctly match the predictor with the criterion variables (those being predicted) for their respective levels of psychometric aggregation in any statistical analysis. For example, higher-order factors such as Super-K should only be used to predict factors that exhibit the same degree of aggregation in their particular hierarchy. The same applies to the K-Factor, which is appropriately matched with the Covitality Factor and the General Factor of Personality (GFP) as hierarchically subordinate indicators of the Super-K Factor, as well as the various lower-order indicators of K-Factor itself (such as Religiosity or Mating Effort) with each other. It is evident from the analyses reported in the critique, in which variables were correlated across quite discrepant levels of aggregation (see Copping, Campbell, and Muncer, 2014, Tables 4, 6, and 7) that any consideration for the correct matching of variables, as required by Brunswik-Symmetry, is wholly absent.

Finally, we address the demographic life event data that are called “the theoretically relevant life history variables” in the critique: age of puberty, age of first sexual encounter, and number of sexual partners. As we noted previously, the relation of these outcome variables to the resource allocation behaviors governed by LH strategy is often ambiguous and confounded by extraneous influences. We need to interject at this point that just because making such connections is difficult does not imply that one should be in any way deterred from exploring them wherever possible; however, one must do so with great caution, without investing excessive faith in certain indicators merely out of historical convention. For example, age of puberty in many primates is partially controlled by nutritional factors that are ultimately limited by the abundance or scarcity of food in the environment and the individual’s social status, as an instrumental factor in obtaining either preferential or dispreferential access to them. Both the age of first sexual encounter and the number of sexual partners that an individual obtains is at least partially limited by the person’s mate quality and the local mate competition in the immediate environment, as might be affected by the local operational sex ratio. Thus, these intermediate outcome factors only partially reflect the process factors that are governed by LH strategy, such as the parametric allocation of bioenergetic and material resources among (1) either somatic or reproductive effort and (2) either mating effort or parental effort. We thus partially agree with the critique in that these dynamic process-outcome relations are worthy of more and better future investigations.

We also partially agree that psychometric measures of LH stand in need of further validation with respect to nonverbal behavior, albeit subject to the enabling conditions or “moderating factors” described above. Our research group has worked hard over the past decade establishing the psychometric properties and exploring the wider nomological net surrounding the construct of human live history strategy. At the present time, our group is working very diligently at establishing construct validity beyond the self-reported questionnaire data that we relied on during the earlier stages of this program of research. In recent years, our research group has been very busy sampling nonverbal behaviors (such as
physical activity, sleep patterns, home-range, risky driving behavior, how people interact in chat rooms and social media, etc.) and correlating them with our better-developed psychometric constructs (formerly based exclusively on verbal self-reports). We do this by various means such as monitoring real-world behaviors using actigraphy, using GPS technology to track daily movements, experimentally placing participants in chat rooms, or observationally counting the frequency and types of activity on social media (e.g., Clacey, Edmunds, and Wolf, 2013; Heany and Wolf, 2013, Sherman, Figueredo, and Funder, 2013; Swanepoel, Thomas, and Wolf, 2013; Swanepoel, Wolf, and Thomas, 2013; Swanepoel, Wolf, Figueredo, and Jacobs, under review, Wolf, Clacey, and Edmunds, 2013; Wolf, Figueredo, and Jacobs, 2013).

We must agree, as psychometricians, with at least some of the arguments that Copping and colleagues present in their critique. Verbal self-report measures may indeed be somewhat biased and may not reflect the actual behavior of participants with complete accuracy. It is possible that self-report instruments only directly capture verbal behavior, but reflect nonverbal behavior to a lesser extent (see Jacobs, Sisco, Hill, Malter, and Figueredo, 2012). The solution, therefore, is to do better psychometrics, not to reject it outright by throwing the proverbial baby out with the bathwater. We nonetheless argue that distinct patterns should emerge among the reports of participants that may at least in part reflect their higher-order latent patterning of cognition, behavior, and reproductive strategy.

Finally, we agree with Copping and colleagues that it would be extraordinary for any LH theorist or empirical researcher to dismiss the importance of critical life events such as “age at first birth” or “number of sexual partners.” The main thrust of our response was directed to their virtually exclusive reliance on single-item verbal self-reports, coupled with their continual tendency to repeatedly refer to these indicators as “objective.” The design of the study used in their critique clearly indicated that three single-item verbal self-reports were being used as the validation criteria for a reasonably well-performing psychometric scale (the HKSS), so what we are making here cannot be construed as a “straw man” argument (see Appendix A for further details on these statistical considerations). That is clearly what is being recommended as preferable by their critique, and it is that particular suggestion with which we are taking serious exception.

Recommendations for Future Research

In this final section, we offer some constructive suggestions for the more successful testing of hypotheses by integrating the biometric and psychometric approaches to the assessment of LH strategies.

Explicit consideration of environmental factors

We have argued that an analysis of biometric indicators is incomplete without an assessment of the environmental conditions in which they occur. As we have explained above, LH theory predicts that an organism’s allocation of bioenergetic and material resources should be calibrated, where appropriate, to suit environmental conditions. Simply measuring a subset of purported biometric LH indicators is inadequate because of the possibly extraneous variance in the environmental conditions that could produce such outcomes. For instance, slow LH strategy is partially characterized by slower growth rates,
later puberty, and fewer offspring. However, the environmental conditions which may produce these outcomes can range from highly abundant, resource-rich environments with high population densities (to increase one’s intraspecific competitiveness) to harsh and barren environments (to preserve one’s bioenergetic and material resources; Brumbach, Figueredo, and Ellis, 2009). These vastly different environments should be expected to produce distinct psychological and behavioral adaptations to cope with their respective challenges. Biometric indicators cannot simply be lumped into one group or another: this is a false dichotomy. The measurement of environmental conditions permits researchers to account for potential confounding factors, such as access to birth control and enhanced levels of nutrition in technologically advanced societies that are expected to influence many of the developmental milestones that are tracked by biometric indicators.

Several environmental factors have been demonstrated to have an effect on some biometric LH indicators, such as father absence on pubertal timing (Ellis and Garber, 2000), and this body of research serves as an excellent starting point. Generally, these variables should assess the relative harshness and unpredictability of the early family environment. More ambitious researchers may examine whether other critical periods (such as adolescence) exist for the influence of these conditions on the calibration of LH strategy.

Although not exhaustive, we offer a sample of theoretically-relevant variables that may be used. Note that these suggestions range from the immediate, familial environment to more distal community or societal environments. Briefly, environmental conditions may include (1) local resource availability and stability, (2) the presence or absence of biological parents, (3) parental conflict, (4) maternal warmth, (5) the presence or absence of genetically-unrelated individuals in the home, (6) the quality of living conditions, (7) the number of changes in residence, (8) local crime rates (especially those impacting extrinsic mortality), (9) indices of socioeconomic inequality such as GINI, and (10) price instability in essential commodities.

**Selection of biometric indicators**

Several biometric indicators should be measured to reduce the attenuating effects of measurement error, and to serve as the basis for factor analyses with which one may examine the associations among biometric factors and psychometric factors (consistent with the Principle of Brunswik-Symmetry). Haphazard selection of indicators is also discouraged; although biometric indicators are expected to be related based on the latent LH construct, initial explorations of these data may be better served by examining “like with like.” In other words, researchers should engage in careful analysis of the potential relationships among biological antecedents and their psychological or behavioral sequelae (or vice versa). For instance, it may be more informative to study the relationships of pubertal timing and fertility rate (biometric indicators of reproductive behaviors) with mating effort, mate value, and sociosexual orientation (psychometric indicators of reproductive behaviors).

Additionally, the measurement of biometric indicators is bolstered by cross-validation using multiple methods of measurement to ensure accuracy in retrospective, cross-sectional studies. Even researchers restricted to self-report data may enjoy the creative exercise in developing multiple items or methods to measure a single biometric parameter. For example, one may consider measuring birth weight using self-report,
maternal-report, hospital records, or archival data. For researchers with more resource availability, longitudinal, observational studies may be feasible and offer even better opportunities to measure relevant biometric parameters in real time; alternatively, hormonal or genetic assays provide even more nuanced developmental data.

Selection of psychometric indicators

As suggested in the previous section, initial empirical research into expanding the nomological net of LH variables may be more informative when the biometric and psychometric variables are selected from the same domain, such as mating effort or parental effort or somatic effort. Following progress in this area, researchers may expand their nomological net to examine resource allocation trade-offs by comparing more than one domain. For example, the study of the relationships between the hypothalamus-pituitary-adrenal axis and hypothalamus-pituitary-gonadal axis would be expected to be reflected in psychological and behavioral trait trade-offs, such as covitality and mating effort.

Finally, researchers who are serious about testing whether psychometric measures of LH map onto biometric measures must include all three of these elements (environmental conditions, biometric indicators, and psychometric indicators) in a single study. Exclusion of one or more of these domains is an inadequate test of hypotheses that seek to expand the nomological net to include both psychometric and biometric variables. However, this area of study is wide open for any who are sincere about developing an accurate and comprehensive measurement model of human LH.

An example application

One application of these concepts is currently underway in the Ethology and Evolutionary Psychology program at the University of Arizona. Two of the co-authors of the present response are examining the structural relations among early life experiences, developmental-biometric markers, and psychometric measures of LH (Black and Cabeza de Baca, in preparation). This study was designed using two novel measures—the Early Environment Questionnaire (Black and Gable, 2012) and the Developmental Markers Questionnaire (Black and Gable, in preparation)—in combination with measures of covitality, mate value, and psychometric measures of LH (Mini-K, HKSS, Mating Effort, Sociosexual Orientation, Time Perspective).

The Early Environment Questionnaire (EEQ) is a self-report measure designed to assess environmental stability and harshness in the first 10 years of life. It focuses on five main domains: 1) Family Economic Resources; 2) Biological Relative Presence; 3) Non-biological Adult Presence; 4) Residential Changes; and 5) Maternal Sensitivity. Although the EEQ focuses on early childhood, a modified version has also been created to assess adolescent environment (Adolescent Environment Questionnaire or AEQ) from age 10–17 in order to explore whether sensitive periods exist in the calibration of LH strategy.

The selection of items for both the EEQ and AEQ was based on two previous works. Griskevicius, Delton, Robertson, and Tybur (2010) measured individual differences in perceived resource-deprivation in childhood and found that disadvantaged participants responded more positively to earlier reproduction after being exposed to a mortality salience prime. Lower childhood SES also predicted earlier age at which participants
wanted to have a first child and the desire to start a family sooner rather than further their career or education (all following the mortality salience prime). Belsky, Schlomer, and Ellis (2012) analyzed longitudinal data from the NICDH Study of Early Child Care and Youth Development. Analyzing the national sample of mothers and children (N = 1364), they found that environmental harshness (defined as income-to-needs ratio in first 5 years of life) and unpredictability (defined as residential changes, paternal transitions, and parental job changes) predicted sexual behavior in 15-year-old adolescents. They also found that the relationship between harshness/unpredictability and adolescent sexual behavior was mediated by maternal depression and sensitivity.

The Developmental Markers Questionnaire (DMQ) is another self-report measure designed to capture several developmental-biometric indicators of life history. It combines novel items about birth weight, pre-mature birth, age at puberty, and age at first sexual experience with a previous measure, called the Pedigree, examining family structure (e.g., family size, interbirth interval, and mother’s age at first birth). It should be noted that this initial exploration into developmental-biometric markers is not exhaustive, and future researchers in this area should generate a sample of possible indicators that covers a broader swathe of developmental history. We also included two measures of general health function (the Rand MOS SF-36) and immune function (Immune Function Survey) to examine, for example, whether physiological states reflect trade-offs between mating effort and somatic effort.

Finally, we included psychometric measures of LH and of psychological and behavioral traits expected to be related to LH. These questionnaires included the Mini-K and HKSS, Mate Value Inventory, Mating Effort Scale, Multidimensional Sociosexual Orientation Inventory, and the Zimbardo Time Perspective Scale. Data collection procedures are still in progress, but we expect to be reporting our results on both student and non-student samples by the time the present work is published.

Acknowledgements: We would like to thank both Dr. Bruce Ellis and Dr. Sally Olderbak, who provided valuable critical feedback as well as references to relevant literature that we had inadvertently overlooked. We would also like to thank Cindy Elizabeth Chavarria-Minera and Mateo Peñaherrera Aguirre, who assisted in the preparation of this paper.

Received 22 May 2014; Revision submitted 17 September 2014; Revision submitted 19 December 2014; Accepted 04 January 2014

References

Abed, R., Mehta, S., Figueredo, A. J., Aldridge S., Balson, H., Meyer, C., and Palmer, R. (2012). Eating disorders and intrasexual competition: Testing an evolutionary hypothesis among young women. The Scientific World Journal, 2012, Article ID 290813, 1-8. doi:10.1100/2012/290813.

Alexander, R. D. (1990a). How did humans evolve? Reflections on the uniquely unique species. Museum of Zoology (Special Publication No. 1). Ann Arbor, MI: The University of Michigan.

Alexander, R. D. (1990b). Epigenetic rules and Darwinian algorithms: The adaptive study of learning and development. Ethology and Sociobiology, 11, 1-63.
Alexander, C. S., Somerfield, M. R., Ensminger, M. E., Johnson, K. E., and Kim, Y. J. (1993). Consistency of adolescents’ self-report of sexual behavior in a longitudinal study. *Journal of Youth and Adolescence, 22*, 455-471.

Belsky, J. (1997). Variation in susceptibility to environmental influence: An evolutionary argument. *Psychological Inquiry, 8*, 182-186.

Belsky, J., Bakermans-Kranenburg, M. J., and Van IJzendoorn, M. H. (2007). For better and for worse: Differential susceptibility to environmental influences. *Current Directions in Psychological Science, 16*, 300-304.

Belsky, J., Scholmer, G. L., and Ellis, B. J. (2012). Beyond cumulative risk: Distinguishing harshness and unpredictability as determinants of parenting and early life history strategy. *Developmental Psychology, 48*, 662-673.

Belsky, J., Steinberg, L., and Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child development, 62*, 647-670.

Bielby, J., Mace, G. M., Bininda-Emonds, O. R. P., Cardillo, M., Gittleman, J. L., Jones, K. E., … Purvis, A. (2007). The fast-slow continuum in mammalian life history: An empirical reevaluation. *The American Naturalist, 169*, 748-757.

Biro, P. A., and Stamps, J. A. (2008). Are animal personality traits linked to life-history productivity? *Trends in Ecology & Evolution, 23*, 361-368.

Black, C. J., and Gable, J. C. (in preparation). The Developmental Markers Questionnaire. Retrieved from [http://www.u.arizona.edu/~cjblack/EEQ.html](http://www.u.arizona.edu/~cjblack/EEQ.html)

Black, C. J., and Cabeza de Baca, T. (in preparation). Using early life experiences to compare developmental and psychometric measures of life history.

Böhnin-Gaese, K., Schuda, M. D., and Helbig, A. J. (2003). Weak phylogenetic effects on ecological niches of Sylvia warblers. *Journal of Evolutionary Biology, 16*, 956-965.

Bouchard, T. J. (1997). Experience producing drive theory: How genes drive experience and shape personality. *Acta Paediatrica, 86*, 60-64.

Bronfenbrenner, U., and Morris, P. A. (2006). The bioecological model of human development. In W. Damon and R. M. Lerner (Eds.), *Handbook of child psychology, Vol. 1: Theoretical models of human development* (6th ed.) (pp. 793-828). New York: John Wiley.

Brumbach, B. H., Figueredo, A. J., and Ellis, B. J. (2009). Effects of harsh and unpredictable environments in adolescence on development of life history strategies: A longitudinal test of an evolutionary model. *Human Nature, 20*, 25–51.

Brumbach, B. H., Walsh, M., and Figueredo, A. J. (2007). Sexual restrictedness in adolescence: A life history perspective. *Acta Psychologica Sinica, 39*, 481-488.

Brunswik, E. (1952). The conceptual framework of psychology (*International encyclopedia of unified science, Volume 1, Number 10*). Chicago: University of Chicago Press.

Buunk, A. P., and Hoben, A. D. (2013). A slow life history is related to a negative attitude towards cousin marriage: A study in three ethnic groups in Mexico. *Evolutionary Psychology, 11*, 442-458.

Buunk, A. P, Pollet, T. V., Klavina, L., Figueredo, A. J., and Dijkstra, P. (2009). Height among women is curvilinearly related to life history strategy. *Evolutionary Psychology, 7*, 545-559.
Cabeza de Baca, T., Figueredo, A. J., and Ellis, B. J. (2012). An evolutionary analysis of variation in parental effort: Determinants and assessment. Parenting: Science and Practice, Special Issue: The Arc of Parenting from Epigenomes to Ethics, 12, 94-104.

Cabeza de Baca, T., Sotomayor-Peterson, M., Smith-Castro, V., and Figueredo, A. J. (2014). Contributions of matrilineal and patrilineal kin alloparental effort to the development of life history strategies and patriarchal values: A cross-cultural life history approach. Journal of Cross Cultural Psychology, 45, 534-554.

Campbell, D. T. (1959). Blind variation and selective survival as a general strategy in knowledge-processes. Self-Organizing Systems, 205-231.

Charnov, E. L. (1993). Life history invariants: Some explorations of symmetry in evolutionary ecology (Vol. 6). Oxford: Oxford University Press.

Clacey, K., Edmunds, C., and Wolf, P. S. A. (2013, March). Sleep, physical activity, affect, and life history strategy: A multi-measure, multi-predictor approach. Presented at the First World Conference on Personality, Stellenbosch, South Africa.

Clark, G. (2008). A farewell to Alms: A brief economic history of the world. Princeton, NJ: Princeton University Press.

Copping, L. T., Campbell, A., and Muncer, S. (2014). Psychometrics and life history strategy: The structure and validity of the high K strategy scale. Evolutionary Psychology, 12, 200-222.

Cronbach, L. J., and Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281.

Crosby, A. W. (2004). Ecological imperialism: The biological expansion of Europe, 900-1900. London, UK: Cambridge University Press.

Darwin, C. (1859). On the origins of species by means of natural selection. London: Murray.

Del Giudice, M. (2009). Sex, attachment, and the development of reproductive strategies. Behavioral and Brain Sciences, 32, 1-21.

Del Giudice, M. (2014). A tower unto heaven: Response to commentaries on “An evolutionary life history framework for psychopathology.” Psychological Inquiry, 25, 394-413.

Del Giudice, M., Ellis, B. J., and Shirtcliff, E. A. (2011). The Adaptive Calibration Model of stress responsivity. Neuroscience and Biobehavioral Reviews, 35, 1562-1592.

Del Giudice, M., Hinnant, J. B., Ellis, B. J., and El-Sheikh, M. (2012). Adaptive patterns of stress responsivity: A preliminary investigation. Developmental Psychology, 48, 775-790.

Dingle, H., and Hegmann, J. P. (Eds.). (1982). Evolution and genetics of life histories. New York: Springer.

Dobson, F. S., and Oli, M. K. (2007). Fast and slow life histories of rodents. In J. O. Wolff and P. W. Sherman (Eds.), Rodent societies: An ecological and evolutionary perspective (pp. 99-105). Chicago: The University of Chicago Press.

Dunkel, C. S., and Decker, M. (2010). Convergent validity of measures of life-history strategy. Personality and Individual Differences, 48, 681-684.

Eckert, C. G. (2000). Contributions of autogamy and geitonogamy to self-fertilization in a mass-flowering, clonal plant. Ecology, 81, 532-542.
Egan, V., Figueredo, A. J., Wolf, P., McBride, K., Sefcek, J. A., Vásquez, G., and Charles, K. (2005). Sensational interests, mating effort, and personality: Evidence for cross-cultural validity. *Journal of Individual Differences, 26*, 11-19.

Ellis, B. J. (2004). Timing of pubertal maturation in girls: An integrated life history approach. *Psychological Bulletin, 130*, 920-958.

Ellis, B. J., Bates, J. E., Dodge, K. A., Fergusson, D. M., John Horwood, L., Pettit, G. S., and Woodward, L. (2003). Does father absence place daughters at special risk for early sexual activity and teenage pregnancy? *Child development, 74*, 801-821.

Ellis, B. J., and Essex, M. J. (2007). Family environments, adrenarche, and sexual maturation: A longitudinal test of a life history model. *Child Development, 78*, 1799-1817.

Ellis, B. J., Figueredo, A. J., Brumbach, B. H., and Schloemer, G. L. (2009). Fundamental dimensions of environmental risk: The impact of harsh versus unpredictable environments on the evolution and development of life history strategies. *Human Nature, 20*, 204-268.

Ellis, B. J., and Garber, J. (2000). Psychosocial antecedents of variation in girls’ pubertal timing: Maternal depression, stepfather presence, and marital and family stress. *Child Development, 71*, 485-501.

Ellis, B. J., McFadyen-Ketchum, S., Dodge, K. A., Pettit, G. S., and Bates, J. E. (1999). Quality of early family relationships and individual differences in the timing of pubertal maturation in girls: A longitudinal test of an evolutionary model. *Journal of Personality and Social Psychology, 77*, 387-401.

Fernandes, H. B. F., Figueredo, A. J., and Woodley of Menie, M. A. (under review). The integration among primate life history traits varies along the fast-slow continuum: A multi-taxa demonstration.

Fernandes, H. B. F., Kruger, D. J., Natividade, J. C., Hutz, C. S., and Kennair, L. E. O. (2012, June). *Postcoital negative emotions: Mental disorder or adaptive mechanism?* Oral presentation at the Human Behavior and Evolution Society, Albuquerque, NM.

Figueredo, A. J. (2007). *The Arizona Life History Battery* [Electronic Version].  
[http://www.u.arizona.edu/~ajf/alhb.html](http://www.u.arizona.edu/~ajf/alhb.html)

Figueredo, A. J., Andrzejczak, D. J., Jones, D. J., Smith-Castro, V., and Montero-Rojas, E. (2011). Reproductive strategy and ethnic conflict: Slow life history as a protective factor against negative ethnocentrism in two contemporary societies. *Journal of Social, Evolutionary, and Cultural Psychology, 5*, 14-31.

Figueredo, A. J., Black, C. J., and Scott, A. G. (2013a). Complementary meta-analytic methods for the quantitative review of research: 1. A theoretical overview. *Journal of Methods and Measurement in the Social Sciences, 4*, 21-45.

Figueredo, A. J., Black, C. J., and Scott, A. G. (2013b). Complementary meta-analytic methods for the quantitative review of research: 2. An extended illustration. *Journal of Methods and Measurement in the Social Sciences, 4*, 46-69.

Figueredo, A. J., Brumbach, B. H., Jones, D. N., Sefcek, J. A., Vásquez, G., and Jacobs, W. J. (2007). Ecological constraints on mating tactics. In G. Geher and G. F. Miller (Eds.), *Mating intelligence: Sex, relationships and the mind’s reproductive system* (pp. 335-361). Mahwah, NJ: Lawrence Erlbaum.
Figueredo, A. J., Cabeza de Baca, T., and Black, C. J. (2014). No matter where you go, there you are: The genetic foundations of temporal stability. *Journal of Methods and Measurement in the Social Sciences, 5*, 76-106.

Figueredo, A. J., Cabeza de Baca, T., and Woodley, M. A. (2013). The measurement of human life history strategy. *Personality and Individual Differences, 55*, 251-255.

Figueredo, A. J., Corral-Verdugo, V., Frías-Armenta, M., Bachar, K. J., White, J., McNeill, P. L., … Castell-Ruiz, I.P. (2001). Blood, solidarity, status, and honor: The sexual balance of power and spousal abuse in Sonora, Mexico. *Evolution and Human Behavior, 22*, 295-328.

Figueredo, A. J., Gladden, P. R., and Hohman, Z. (2012). The evolutionary psychology of criminal behavior. In S.C. Roberts, (Ed.), *Applied evolutionary psychology* (pp. 201-221). New York: Oxford University Press.

Figueredo, A. J., Hammond, K. R., and McKiernan, E. C. (2006). A Brunswikian evolutionary developmental theory of preparedness and plasticity. *Intelligence, 34*, 211-227.

Figueredo, A. J., McKnight, P. E., McKnight, K. M., and Sidani, S. (2000). Multivariate modeling of missing data within and across assessment waves. *Addiction, 95 (Supplement 3)*, S361-S380.

Figueredo, A. J., and Rushton, J. P. (2009). Evidence for shared genetic dominance between the general factor of personality, mental and physical health, and life history traits. *Twin Research and Human Genetics, 12*, 555-563.

Figueredo, A. J., Tal, I. R., McNeill, P., Guillén, A. (2004). Farmers, herders, and fishers: The ecology of revenge. *Evolution and Human Behavior, 25*, 336-353.

Figueredo, A. J., Vásquez, G., Brumbach, B. H., and 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., and Schneider, S. M. R. (2007). The K-factor, covitality, and personality: A psychometric test of life history theory. *Human Nature, 18*, 47-73.

Figueredo, A. J., Vásquez, G., Brumbach, B. H., Schneider, S. M. R., Sefcek, J. A., Tal, I. R., … Jacobs, W. J. (2006). Consilience and life history theory: From genes to brain to reproductive strategy. *Developmental Review, 26*, 243-275.

Figueredo, A. J., and Wolf, P. S. A. (2009). Assortative pairing and life history strategy: A cross-cultural study. *Human Nature, 20*, 317-330.

Figueredo, A. J., Wolf, P. S. A., Olderbak, S. G., Gladden, P. R., Fernandes, H. B. F., Wenner, C., … Rushton, J. P. (2014). The psychometric assessment of human life history strategy: A meta-analytic construct validation. *Evolutionary Behavioral Sciences, 8*, 148-185.

Figueredo, A. J., Woodley, M. A., Brown, S. D., and Ross, K. C. (2013). Multiple successful tests of the Strategic Differentiation-Integration Effort (SD-IE) hypothesis. *Journal of Social, Evolutionary, and Cultural Psychology, 7*(4), 361.

Figueredo, A. J., Woodley of Menie, M. A., and Jacobs, W. J. (in press). The evolutionary psychology of the general factor of personality: A hierarchical life history model. In D. Buss (Ed.), *Handbook of evolutionary psychology (2nd ed.)*. Hoboken, NJ: Wiley.

Flere, S., and Lavrič, M. (2008). On the validity of cross-cultural social studies using student samples. *Field Methods, 20*, 399-412.
Flinn, M. V. (2006). Evolution and ontogeny of stress response to social challenges in the human child. *Developmental Review, 26*, 138-174.

Flinn, M. V., Geary, D. C., and Ward, C. V. (2005). Ecological dominance, social competition, and coalitionary arms races: Why humans evolved extraordinary intelligence. *Evolution and Human Behavior, 26*, 10-46.

Frías-Armenta, M., Valdez, P., Nava, M. G., Borrani, J., Figueredo, A. J., Contreras, M., … Ríos, D. (2005). *Funciones ejecutivas y delincuencia juvenil: datos preliminares*. In M. Frías-Armenta and V. Corral-Verdugo (Eds.), *Niñez, adolescencia, y problemas sociales* (pp. 267-278). México: CONACYT-UniSon.

Frías-Armenta, M., Valdez-Ramirez, P., Nava-Cruz, M. G., Figueredo, A. J., and Corral-Verdugo, V. (2010). Executive functions, attention and juvenile delinquency. In M. Frías-Armenta and V. Corral-Verdugo (Eds.), *Bio-psycho-social perspectives on interpersonal violence* (pp. 29-50). Hauppauge, NY: Nova Science Publishers.

Gaillard, J. M., Yoccoz, N. G., Lebreton, J. D., Bonenfant, C., Devillard, S., Loison, A., ... Allaine, D. (2005). Generation time: A reliable metric to measure life-history variation among mammalian populations. *The American Naturalist, 166*, 119-123.

Gallup (2009). *Gallup-Healthways Well-Being™ Index: Methodology report for indexes*. Retrieved from [http://wbi.meyouhealth.com/files/GallupHealthwaysWBI-Methodology.pdf](http://wbi.meyouhealth.com/files/GallupHealthwaysWBI-Methodology.pdf)

Galton, F. (1869). *Hereditary genius: An inquiry into its laws and consequences*. London: Macmillan.

Garcia, R. A., Cabeza de Baca, T., Sotomayor-Peterson, M., Smith-Castro, V., and Figueredo, A. J. (in preparation). Measures of domain-specific resource allocations in life history strategy: Indicators of a latent common factor or ordered developmental sequence?

Gaxiola-Romero, J., and Frías-Armenta, M. (2008). Un modelo ecológico de factores protectores del abuso infantil: Un estudio con madres Mexicanas. *Medio Ambiente y Comportamiento Humano, 9*, 13-21.

Giosan, C. (2006). High-K strategy scale: A measure of the high-K independent criterion of fitness. *Evolutionary Psychology, 4*, 394-405.

Gladden, P. R., Figueredo, A. J., Andrejzak, D. J., Jones, D. N., and Smith-Castro, V. (2013). Reproductive strategy and sexual conflict: Slow life history strategy inhibits negative androcentrism. *Journal of Methods and Measurement in the Social Sciences, 4*, 48-71.

Gladden, P. R., Sisco, M., and Figueredo, A. J. (2008). Sexual coercion and life-history strategy. *Evolution and Human Behavior, 29*, 319-326.

Griskevicius, V., Delton, A. W., Robertson, T. E., and Tybur, J. M. (2010). Environmental contingency in life history strategies: The influence of mortality and socioeconomic status on reproductive timing. *Journal of Personality and Social Psychology, 100*, 241-254.

Hamilton, D. T., and Morris, M. (2010). Consistency of self-reported sexual behavior in surveys. *Archives of Sexual Behavior, 39*, 842–860.

Harvey, P. H., and Purvis, A. (1999). Understanding the ecological and evolutionary reasons for life history variation: Mammals as a case study. In J. McGlade (Ed.), *Advanced ecological theory: Principles and applications* (pp. 232-247). Oxford: Blackwell Science.
Heany, S., and Wolf, P. S. A. (2013, March). The effects of mortality salience on life history strategy and religiosity: A terror management study with an evolutionary perspective. Presented at The 1st World Conference on Personality, Stellenbosch, South Africa.

Henrich, J., Heine, S. J., and Norenzayan, A. (2010). The weirdest people in the world? Behavioral and Brain Sciences, 33, 61-135.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30, 179-185.

Hrdy, S. B. (1999). Mother nature: A History of mothers, infants, and natural selection. New York: Pantheon.

Irons, W. (1990). Adaptively relevant environments versus the environment of evolutionary adaptedness. Evolutionary Anthropology, 6, 194-204.

Jacobs, W. J., Sisco, M., Hill, D., Malter, F., and Figueredo, A. J. (2012). On the practice of theory-based evaluation: Information, norms, and adherence. Evaluation and Program Planning, 35, 354–369.

Jeschke, J. M., and Kokko, H. (2009). The roles of body size and phylogeny in fast and slow life histories. Evolutionary Ecology, 23, 867-878.

Johansson, F. (2000). The slow-fast life style characteristics in a suite of six species of odonate larvae. Freshwater Biology, 43, 149-159.

Johnson, W., Turkheimer, E., Gottesman, I. I., and Bouchard, T. J. Jr. (2010). Beyond heritability: Twin studies in behavioral research. Current Directions in Psychological Science, 18, 217-220.

Jonason, P. K., Li, N. P., and Czarna, A. Z. (2011). Quick and dirty: Some psychosocial costs associated with dark triad in three countries. Evolutionary Psychology, 11, 172-185.

Kamilar, J. M., and Cooper, N. (2013). Phylogenetic signal in primate behaviour, ecology and life history. Philosophical Transactions of the Royal Society B: Biological Sciences, 368, 20120341.

Kaplan, H. S., and Gangestad, S. W. (2005). Life history theory and evolutionary psychology. In D. M. Buss (Ed.), Handbook of evolutionary psychology (pp. 68-95). New York: Wiley.

Kline, R. B. (2011). Principles and practice of structural equation modelling (3rd ed.). New York: Guilford.

Ledesma, R. D., and Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. Practical Assessment, Research and Evaluation, 12, 1-11.

Little, T. D., Cunningham, W. A., Shahar, G., and Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question and weighing the merits. Structural Equation Modeling, 9, 151-173.

Mayr, E. (1982). The growth of biological thought: Diversity, evolution, and inheritance. Cambridge, MA: Belknap Press of Harvard University Press.

McKnight, P. E., McKnight, K. M., Sidani, S., and Figueredo, A. J. (2007). Missing data: A gentle introduction. New York: Guilford.

Meaney, M. J. (2010). Epigenetics and the biological definition of gene×environment interactions. Child development, 81, 41-79.
Meston, C. M., Heiman, J. R., Trapnell, P. D., and Paulhus, D. L. (1998). Socially desirable responding and sexuality self-reports. *The Journal of Sex Research, 35*, 148-157.

Must, A., Phillips, S. M., Naumova, E. N., Blum, M., Harris, S., Dawson-Hughes, B., and Rand, W. M. (2002). Recall of early menstrual history and menarcheal body size: After 30 years, how well do women remember? *American Journal of Epidemiology, 155*, 672-679.

Nunnally, J., and Bernstein, I. (1994) *Psychometric theory* (3rd ed.). New York: McGraw Hill.

Oli, M. K. (2004). The fast-slow continuum and mammalian life-history patterns: An empirical evaluation. *Basic and Applied Ecology, 5*, 449-463.

Pennanen, T., Perkiömäki, J., Kiikkilä, O., Vanhala, P., Neuvonen, S., and Fritze, H. (1998). Prolonged, simulated acid rain and heavy metal deposition: Separated and combined effects on forest soil microbial community structure. *FEMS Microbiology Ecology, 27*, 291-300.

Petrinovich, L. (1979). Probabilistic functionalism: A conception of research method. *American Psychologist, 34*, 373-390.

Pianka, E. R. (1970). On r- and K-selection. *American Naturalist, 104*, 592-597.

Promislow, D. E., and Harvey, P. H. (1990). Living fast and dying young: A comparative analysis of life-history variation among mammals. *Journal of Zoology, 220*, 417-437.

Read, A. F., and Harvey, P. H. (1989). Life history differences among the eutherian radiations. *Journal of Zoology, 219*, 329-353.

Reznick, D., Bryant, M. J., and Bashey, F. (2002). r-and K-selection revisited: the role of population regulation in life-history evolution. *Ecology, 83*(6), 1509-1520.

Roff, D. A. (2002). *Life history evolution, Vol. 7*. Sunderland: Sinauer Associates.

Rowe, D. C. (1994). *The limits of family influence: Genes, experience, and behavior*. Guilford Press.

Rushton, J. P. (1985). Differential K theory: The sociobiology of individual and group differences. *Personality and Individual Differences, 6*, 441-452.

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.

Schulte-Hostedde, A. I., Millar, J. S., and Hickling, G. J. (2001). Evaluating body condition in small mammals. *Canadian Journal of Zoology, 79*, 1021-1029.

Sherman, R. A., Figueredo, A. J., and Funder, D. C. (2013). The behavioral correlates of overall and distinctive life history strategy. *Journal of Personality and Social Psychology, 105*, 873-888.

Simpson, J. A., Griskevicius, V., Kuo, S. I., Sung, S., and Collins, W. A. (2012). Evolution, stress, and sensitive periods: The influence of unpredictability in early versus late childhood on sex and risky behavior. *Developmental Psychology, 48*, 674.

Smith, T. W. (1992). Discrepancies between men and women in reporting number of sexual partners: A summary from four countries. *Biodemography and Social Biology, 39*, 203-211.

Sotomayor-Peterson, M., Cabeza de Baca, T., Figueredo, A. J., and Smith-Castro, V. (2013). Shared parenting, parental effort, and life history strategy: A cross-cultural comparison. *Journal of Cross-Cultural Psychology, 44*, 620-639.
Sotomayor-Peterson, M., Figueredo, A. J., Hendrickson-Christensen, D., and Taylor, A. R. (2012). Couples’ cultural values, shared parenting, and family emotional climate within Mexican-American families. *Family Process, 51*, 218-233.

Stearns, S. C. (1992). *The evolution of life histories*, Vol. 248. Oxford: Oxford University Press.

Stearns, S. C., and Koella, J. C. (1986). The evolution of phenotypic plasticity in life-history traits: Predictions of reaction norms for age and size at maturity. *Evolution, 40*, 893-913.

Stevens, J. R. (2014). Evolutionary pressures on primate intertemporal choice. *Proceedings of the Royal Society B: Biological Sciences, 281*, 20140499.

Sullivan, J. L., and Feldman, S. (1979). Multiple indicators: An introduction. *Quantitative applications in the social sciences, no. 07-015*. Beverly Hills, CA: Sage Publications.

Swanepoel, T., Thomas, K.G. F., and Wolf, P. S. A. (2013, March). *Sexual strategies among young adults in the virtual world*. Paper presented at The 1st World Conference on Personality, Stellenbosch, South Africa.

Swanepoel, T., Wolf, P. S. A., and Thomas, K. G. F. (2013, July). *Life history status predicts the form and success of sexual strategies in college-age men and women*. Paper presented at the Human Behavior and Evolution Society, Miami, Florida.

Swanepoel, T., Wolf, P. S. A., Jacobs, W. J., and Thomas, K. G. F. (under review). *Life history status predicts the form and success of sexual strategies in college-age men and women*.

Tal, I. R., Hill, D., Figueredo, A. J., Frías-Armenta, M., and Corral-Verdugo, V. (2006). An evolutionary approach to explaining water conservation. *Medio Ambiente y Comportamiento Humano, 7*, 7-27.

Taylor, A, West, S., and Aiken, L (2005, August). *Effects on power of coarsely categorizing an outcome variable*. Presented at the American Psychological Association, Washington, DC.

Tifferet, S., Agrest, S., and Shlomo Benisti, A. (2011). Problem gambling: An outcome of a life history strategy. *International Gambling Studies, 11*, 253-262.

Tooby, J., and Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology, 11*, 375-424.

Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell, (Ed.), *Sexual selection and the descent of man* (pp. 136-179). London: Heinemann.

van der Linden, D., Figueredo, A. J., de Leeuw, R. N. H., Scholte, R. H. J., and Engels, R. C. M. E. (2012). The general factor of personality (GFP) and parental support: Testing a prediction from life history theory. *Evolution and Human Behavior, 33*, 537-546.

van Schaik, C. P., and Isler, K. (2012). Life-history evolution. In J. C. Mitani, J. Call, P. M. Kappeler, R. Palombit, and J. B. Silk (Eds.), *The evolution of primate societies* (pp. 220-244). Chicago: University of Chicago Press.

Wemple, S. F. (1981). *Women in Frankish society: Marriage and the cloister, 500 to 900*. Philadelphia: University of Pennsylvania Press.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. New York: Oxford University Press.
Wolf, P. S. A., Clacey, K., and Edmunds, C. (2013, July). *Sleep, physical activity, and life history strategy: A multiple-predictor study*. Presented at the Human Behavior and Evolution Society, Miami, Florida.

Wolf, P. S. A., Figueredo, A. J., and Jacobs, W. J. (2013). Global positioning system technology (GPS) for psychological research: A test of convergent and nomological validity. *Frontiers in Psychology, 4*, 315.

Woodley, M. A., Figueredo, A. J., Brown, S. D., and Ross, K. C. (2013). Four successful tests of the cognitive differentiation-integration effort hypothesis. *Intelligence, 41*, 832-842.

Woodley of Menie, M. A., and Madison, G. (2015). The association between $g$ and $K$ in a sample of 4246 Swedish twins: A behavior genetic analysis. *Personality and Individual Differences, 74*, 270-274.
Appendix: Statistical Concerns and Clarifications

The internal structure of the HKSS

One recommendation offered in the critique was the elimination of HKSS items that might be omitted by certain respondents pertaining to children, marriage, and living with a partner. This suggestion fails to take into account the fact that these characteristics are theoretically related to life history strategy. Instead of calling for their elimination, it would have been more helpful to examine whether individuals who were characterized as faster life history by the HKSS were more likely to omit these questions than others. In cases such as these, not responding to an item because it “does not apply” is itself a source of valuable information. As statisticians and methodologists have noted for years (most recently in Figueredo, McKnight, McKnight, and Sidani, 2000; McKnight, McKnight, Sidani, and Figueredo, 2007), patterns of missingness can yield important insights into the individual characteristics of participants who either do or do not answer each specific question. Another possibility is that many respondents do not answer those specific questions about marriage because they are too young to marry or about children because they as yet have none to report. Thus, although these items might have been better worded, their inclusion is not a serious flaw in the measure. We note that the mean age of the sample reported on in the critique is only 39 years old, which leaves some latitude for a non-negligible number of the as yet single and childless.

Another attempt to improve upon the measure undertaken in the critique is an examination of the internal structure of the HKSS. We believe that this exercise is inadvisable because the HKSS is a relatively short 26-item scale, and it is remarkable that it generally has as high a reliability coefficient as it actually does. Nevertheless, the critique splits the already-short scale into four even shorter subscales originally “hypothesized” by Giosan (2006), as represented in the critique. Each of these reconstructed subscales is comprised of 3–4 items each, which weakens each one psychometrically, and it is hard to imagine that almost any existing short form could survive being decomposed in that manner. Our interpretation of Giosan’s (2006) breakdown of the HKSS items into these specific groupings was not to recommend their use as separate subscales, but instead to clarify the nature of the four content domains sampled by the HKSS.

The decomposition is especially psychometrically problematic when coupled with a confirmatory factor analysis using these 3–4 item subscales as indicators. This is because confirmatory factor analysis penalizes the inclusion of low-reliability indicators in terms of goodness-of-fit. Given that the number of items in any scale constitutes an intrinsic part of the computational algorithm for Cronbach’s alpha, such tiny subscales will necessarily tend to show very low internal consistency reliabilities. That being the case, this procedure is therefore virtually “burying” the factor structure in measurement error (which is the complement of the reliability coefficient). Even so, when one examines the factor loadings of the HKSS factor on the subscales, they range from .55 to .95, whether using the original Giosan specification or their own respecification of the subscales. These are factor loadings that throughout the rest of psychometrics would have been considered to be coefficients of fairly respectable magnitude as indicators of convergent validity. The only reason given for its rejection in the critique was on the basis of its lower indices of goodness-of-fit in relation to their preferred model, which was of four correlated “factors” (meaning
subscales) that did not converge upon a higher-order construct. The problems with these goodness-of-fit indicators as applied to the design of their models render the analyses offered in the critique psychometrically invalid.

The hierarchical latent structure of Life History Strategy

One exception to the general pattern of results supporting a higher-order life history factor was an inappropriate analysis (Bielby et al., 2007) that (1) applied a Varimax factor rotation, specifically designed to produce orthogonal (uncorrelated) factors; (2) used samples with as few observations as \( n = 40 \) for a multivariate analysis, which is highly inadvisable; (3) controlled statistically for body weight, which unnecessarily eliminates the common factor variance associated with one of its major indicators, and thus eliminates much of the variance due to genuine life history trade-offs (Jesche and Kokko, 2009; Roff, 2002); and (4) controlled statistically for the effects of phylogeny, using the method of independent phylogenetic contrasts, which assumes that the phylogenetic signal (a measure of similarity between closely related species compared to distantly related species) is at a maximum. The latter method thus artificially reduces the magnitudes of the covariation among traits, as life history traits have been demonstrated to have only modest levels of phylogenetic signal (e.g., Kamilar and Cooper, 2013).

The purported objectivity of biometric indicators

A further methodological concern that we need to raise is the repeated reference to these variables as “objective measures” against which psychometric assessments of life history need to be validated. Granted, these are some of the fundamental dimensions on which most of the original research on life history strategy in nonhuman animals and plants was based. However, in evolutionary ecological studies, these dimensions are usually directly observed or directly measured in the field. Several authors of the present response to the critique were originally trained in the biological sciences, such as comparative behavioral and molecular ecology, and spent many years studying nonhuman animal and plant reproductive behaviors by direct observation (e.g., actual bioenergetic allocations, gene expression, growth, mating, and reproduction). We therefore know how this is done within the natural sciences.

The demographic life event data that are being represented as “objective” in the critique, however, are evidently based exclusively on self-reports, not upon direct observations. No institutional review board for research involving human subjects would provide ethical approval for direct observations of any of the following: (1) the first menstrual discharge of each participant, (2) the first sexual experience of each participant, or (3) the actual number of sexual partners with which each participant had coitus. Assuming that such direct observations were not conducted, we fail to understand why these particular self-report measures are to be considered “objective.” Granted, these items explicitly refer to life events which must have objectively either happened or not happened in the material world, but so do most of the items in psychometric inventories such as the ALHB (http://www.u.arizona.edu/~ajf/alhb.html). These self-report data, however, cannot be considered completely objective, whether they are being applied by well-trained and qualified demographers or by professional psychometricians.
The reason is that items like these are clearly subject to extraneous psychological influences, such as self-presentation biases, such as self-deceptive enhancement and impression management (for examples in sexuality research, see Meston, Heiman, Trapnell, and Paulhus, 1998). When you ask women retrospectively about the age at which they first reached puberty, their recall lacks sufficient levels of precision. In one longitudinal study (Must et al., 2002), for example, the precision with which adult women were able to retrospectively identify their age at menarche up to 30 years later was plus or minus about one half year for 55% of the sample and about one full year for 79% of the sample (inclusive of the 55%). These degrees of accuracy are celebrated as good news in the research report on the study, but it would be difficult to detect the effects of any individual differences in life history with that degree of assessment error in the estimate (although one could make aggregated group comparisons with this margin of error with the use of sufficiently large sample sizes).

Test–retest reliabilities have been estimated in several long-term prospective studies in which women’s self-reported age at menarche was first obtained in adolescence and then again 17 to 37 years later (Ellis, 2004); correlations across these two measurement periods have been consistently high, ranging from .67 to .79, with comparable data not available for men. Correlations within this range, however, might indeed have considerable errors of prediction associated with them, in spite of their being a systematic association between pretest and posttest scores. We believe that the margin for error plus or minus one half or one full year (reported in Must et al., 2002) is more intuitively indicative of just how imprecise such retrospective reports of age of menarche might be. The correlation coefficient tells you how well you can estimate aggregate group statistics in a predictive model; the “error bars” tell you just how off they are likely to be in practice at the individual level.

Moreover, when you ask men about the same thing, their recall ranges from nonexistent to unclear due to the relative lack of overt external signs shown by young men as compared with young women, a problem that is also mentioned in the critique itself. Similarly, when you ask men and women to verbally report either their age at first intercourse or their lifetime number of sexual partners, even when guaranteeing complete confidentiality and anonymity, you are not very likely to get an unbiased answer. For example, based on mainstream evolutionary psychological theory, we expect that women should generally overestimate and men should generally underestimate their age at first intercourse (Alexander, Somerfield, Ensminger, Johnson, and Kim, 1993), whereas women should generally underestimate and men should generally overestimate their number of sexual contacts (Smith, 1992), for the relatively obvious reasons of the sexually-dimorphic value judgments of social and sexual desirability ascribed to those numbers in virtually any human culture. Hamilton and Morris (2010) review of a multitude of tests that assessed the reliability of self-report measures of sexual behavior, pointing out that “the test-retest correlations have been as low as .3 and as high as .9 across these various populations” (p. 843). In view of this finding, they are very surprised to find that their own study found a high consistency for their particular sexuality-related subject. In spite of these occasionally high test-retest consistencies, our point was that there are consistent biases in self-reported sexual behavior data, such as women under-reporting the number of partners they have had (Smith, 1992). This means that women can provide highly consistent verbal responses
across time (producing high test-retest reliabilities) but those responses might nonetheless be biased in a variety of ways and not accurately represent what happened in the real world.

Even when couples purport to be in a monogamous romantic relationship, male partners generally tend to separately report having sex more frequently than do their female partners; such discrepancies have also been observed in the reporting by men and women of male perpetration as compared to female victimization in the frequencies of spousal abuse (see Figueredo et al., 2001). Unless the couple is not completely monogamous, there is no way to reconcile those discrepant reports from men and women as simultaneously representing objective facts. The claim repeatedly made in the critique to the effect that their demographic life event data are objective would therefore be considered unacceptable in psychometrics.

Of course, not all single item measures are created equal. Some items, such as asking a person’s chronological age, can be a reliable and valid item, whereas other items that ask about the same person’s opinions on a Likert scale may be less reliable and valid. In general, questions about actual behaviors or events (colloquially called behavioroid items) tend to be more reliable, due to the objective nature of behaviors and events. Even when asking participants about seemingly objective events, the way you ask the questions and the response format provided may impact reliability and validity by introducing both random errors and systematic biases. When relying on single items, every attempt should be made to ensure that the item is as well-written and thought out as possible. We contend therefore that the authors of the critique could have done a better job of designing their items.

For example, the authors of the critique even speculated that the puberty question would be more unreliable for males because male signs of puberty are often less memorable. We agree, especially as they asked the participants how old they were when they reached puberty without defining what “reaching puberty” means. They left it up to the participants to define the concept, and because of this, the participants might be reporting on differing life events that are indicators of pubertal development but do not happen simultaneously, even if they had good memories of events that happened on average 20 to 30 years ago.

Given the likely differential reliability of the measure, where females are more reliable at estimating their age at puberty than males, it is not surprising that there were positive correlations between age of puberty and the HKSS total scores and three of the four subscales in females and that these statistical patterns where not found in males. Reliability is an important component when it comes to a study’s ability to detect an effect. Each of the items should have been worked on (and perhaps pilot-tested on an independent sample) before they were presented to participants, and this item is not even the most problematic one.

Knowing nothing else, besides what is presented in the critique, the number of sexual partners question is probably the most problematic item of the lot. They used a 7-item scale to capture a continuum ranging from zero to infinity. The authors then took this ordinal scale and decided to correct for reproductive lifespan. To be able to do so, they divided the lower bound of each of the ordinal categories (i.e., 0, 1, 2, 11, 21, 51, and 100) by reproductive lifespan, which was calculated by subtracting the participants’ age of puberty from the participants’ chronological age. Setting aside whether or not it is advisable to use a nominal variable to create an index that is composed of a ratio, the
procedure the authors chose to use should raise red flags on issues related to statistical power, error, and bias.

The authors of the critique also reported that the mean number of sexual partners for the sample was .37 and the standard deviation was .76. As nobody should be able to score negative numbers on that scale, these parameters indicate a positively skewed distribution. We bring this up because coarsely categorizing a continuous dependent variable has negative impacts on statistical power, especially when the variable of interest is skewed (Taylor, West, and Aiken, 2005). The authors may argue that their dependent variable was more continuous than the raw score implies. Although this has to be true, the problem remains that a theoretically important component of their index was coarsely categorized, which will impact the statistical power necessary to detect an effect (if there was one), and that the source of this extra variance is expected to cause other problems that we discuss below.

The procedure they decided to use probably distorts reality in several ways. First, because they took the lower bound of each of the ordinal categories and then corrected for reproductive age, it means that two people with the same length of reproductive life who in reality had 2 and 10 reproductive partners, respectively, would have had the same estimate of lifetime reproductive partners and the same estimate for yearly partners corrected for reproductive lifespan. That assigned score would be 2. If their reproductive lives are both 10 years long, the number of partners per year per participant would be .2, even though one of them had, on average, 1 reproductive partner per reproductive year. This procedure will do three things: (1) because the lower bound of each category was used, the measure will systematically underestimate number of sexual partners; (2) because the numerator only has seven categories and the measures in the denominator being subtracted from each other are both measured on continuous interval scales, most of the variance in scores in their index is going to be related to the denominator and not the more important numerator; and (3) because the categories’ range of possible answers systematically increase the more sexual partners a participant reports, the error in measurement is going to increase systematically as more reproductive partners are reported.

This procedure will do more than impact statistical power. As the denominator in the index used in the critique probably contributes most of the variance in that measure, it means that the index is likely to reflect chronological age more than number of sexual contacts achieved during that time. In other words, if one correlates this index with its components, the highest correlations would most likely be with the age variable. Therefore the reported correlation of this index with the HKSS is more of a correlation with age than number of sexual partners per year. This goes beyond the restriction of range problem. The combination of two variables at two levels of measurement can have strange consequences.

This scenario is analogous to one that might be encountered in a repeated measures design. Consider a hypothetical 20 year old male, who reached puberty 5 years ago, had only one sexual partner, and then had an additional sexual partner every subsequent year. Using this procedure, that person’s average number of partners per reproductive year will have started out at .2, and continued to decrease every year until he turned 30 and graduated to the next age category. By the time he turned 29, he would have been assigned an average of .14 sexual partners, when a more accurate measure should total about .71. Translated to the population level, this implies that unless a respondent’s actual number of sexual partners lies right at the bottom of one of the arbitrary age categories, there will be a
systematic underestimation of number of sexual partners and that underestimate gets worse for older participants. Furthermore, as the number of sexual partners reported increases, the measurement procedure becomes increasingly coarse, which inevitably results in increasingly severe degrees of underestimation. This generates two identifiable sources of bias that are correlated with each other, because older participants should generally have had more lifetime sexual partners.

We consider that taking the least developed psychometric life history short form, sampling British middle-class adults, and correlating the scores to admittedly unreliable measures may not be best way to make a persuasive argument against the psychometric approach to assessing human life history strategy. Recall that failing to find a statistical relationship can be attributable to several different possible causes. When the measures used are questionable, two options come to mind: (1) there is no real relationship in the population, and (2) the study does not have the requisite statistical power to detect an effect.

How then can the entire field of psychometrics be founded on such questionable data? Simply put, psychometric methods involve taking great precautions in the interpretation of self-report data. These sophisticated methods are too complex to detail in this response, but there are many references on the subject readily available (e.g., Nunnally and Bernstein, 1994). The one that we will emphasize here, however, is the Psychometric Principle of Aggregation (see Sullivan and Feldman, 1979). Simply put, it means that aggregating across a large number of items substantially reduces measurement error. If the items are also methodologically heterogeneous, then it also may reduce measurement bias; otherwise, it is by no means guaranteed to do so. Nevertheless, it is well-established that the psychometric reliability of any single item is generally quite low, and only when composited into scales do we achieve acceptable levels of reliability. The low reliabilities of single items therefore introduce a great deal of random measurement errors into any predictive equation and thus function to systematically attenuate any correlation coefficients, regression weights, or structural parameters estimated from those data. The separate use of single items is therefore generally considered highly inadvisable in psychometrics. For example, Nunnally and Bernstein (1994) have argued specifically against using individual items as indicators. More recently, Kline (2011) has noted that:

Another problem is that item-level data tend to be “noisy.” Specifically, people’s responses to individual items may be unstable, so item reliabilities can be low. Items in exploratory factor analysis (EFA) often have relatively high secondary loadings (e.g., about .30) on factors other than the one on which they have primary loadings (e.g., > .50). Secondary loadings in EFA often account for relatively high proportions of the variance, so constraining them to zero in CFA may be too conservative. Consequently, the more restrictive CFA model may not fit the data. This is one reason the specification of a CFA model based on EFA outcomes and analyzed with the same data may lead to the rejection of the CFA model (van Prooijen and van der Kloot, 2001). That is, CFA does not generally “confirm” the results of EFA. (p. 244)

For these reasons, many psychometricians recommend the use of easily-created parcels, as opposed to single items, as manifest variables (an approach advocated by Little,
Cunningham, Shahar, and Widaman, 2002). As Kline (2011) also noted, “An alternative to analyzing items as indicators with special estimators is to analyze parcels with a normal theory method, such as ML” (p. 244). Such parcels are created from 3–4 items each, and would therefore be analogous to the composition of the HKSS “subscales” under scrutiny. The analysis presented in the critique could easily have been done this way simply by using the subscales (qua “parcels”) as the manifest indicators in the confirmatory factor models. Instead, the analysis reported in the critique loads the subscales (qua “factors”) on the individual constituent items, and the fit of all the models is severely compromised by the resulting errors of measurement, which serve to inflate the residuals on which the goodness-of-fit indices are based.

The analyses of the HKSS reported in the present critique rely primarily upon three separate self-reports of life events: age of puberty, age of first sexual encounter, and number of sexual partners. These three self-report items are not composited into a single scale and the magnitudes of the bivariate correlations among them are not reported. We wonder what the internal consistency might be for that scale, if they were aggregated with each other into their own “parcel.” We expect that it would not be very high, based on the reasoning previously applied to the decomposed “subscales” of the HKSS. In spite of all this, the critique uses these three self-report items as the ultimate standard against which several much better-developed and better-validated multi-item and sometimes multi-factor scales are to be evaluated. We find this recommendation quite surprising in view of the nearly universal rejection of this general approach in psychometric theory (e.g., Kline, 2011; Nunnally and Bernstein, 1994).