



Advancing the Psychometric Study of Human Life History Indicators

K Does Not Measure Life History Speed, but Theory and Evidence Suggest It Deserves Further Attention

George B. Richardson¹ · Nathan McGee¹ · Lee T. Copping²

Accepted: 20 April 2021 / Published online: 28 May 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

In this article we attend to recent critiques of psychometric applications of life history (LH) theory to variance among humans and develop theory to advance the study of latent LH constructs. We then reanalyze data ($n=4,244$) previously examined by Richardson et al. (*Evolutionary Psychology*, 15(1), 2017, <https://doi.org/10.1177/1474704916666840>) to determine whether (a) previously reported evidence of multidimensionality is robust to the modeling approach employed and (b) the structure of LH indicators is invariant by sex. Findings provide further evidence that a single LH dimension is implausible and that researchers should cease interpreting K-factor scores as empirical proxies for LH speed. In contrast to the original study, we detected a small inverse correlation between mating competition and Super-K that is consistent with a trade-off. Tests of measurement invariance across the sexes revealed evidence of metric invariance (i.e., equivalence of factor loadings), consistent with the theory that K is a proximate cause of its indicators; however, evidence of partial scalar invariance suggests use of scores likely introduces bias when the sexes are compared. We discuss limitations and identify approaches that researchers may use to further evaluate the validity of the K-factor and other applications of LH to human variation.

Keywords Life history theory · Life history strategy · Structure of life history indicators · Sex differences · Exploratory structural equation modeling

✉ George B. Richardson
george.richardson@uc.edu

¹ University of Cincinnati, Cincinnati, OH, USA

² Teesside University, Middlesbrough, UK

Research on individual differences in human life history (LH) strategy has proliferated (for reviews, see Black et al., 2017; Ellis et al., 2009) following seminal applications of LH theory to human development (Belsky et al., 1991; Draper & Harpending, 1982). During the past couple of decades, psychometric research extended the application of LH theory to a broad suite of traits (Black et al., 2017; Figueredo et al., 2006). Most psychometric applications of LH theory have assumed that a single latent variable adequately captures the covariance among an array of indicators, justifying the use of a single score as an empirical proxy for LH speed. Consistent with this, a second-order K-factor subsumed indicators including planning and control, social contact and support, attachment, religiosity, and altruism (Figueredo et al., 2004, 2005, 2007). Moreover, a third-order Super-K factor subsumed the second-order K-factor, Covitality (i.e., health and mental health), and the General Factor of Personality (GFP; Figueredo et al., 2004, 2005, 2007; Olderbak et al., 2014). Recently, however, several theoretical and empirical challenges to the measurement of human LH strategy via a single score have emerged.

Psychometric Critiques

In the first psychometric critique to focus on second-order LH measures, Copping et al. (2014) pointed out that the internal structure of the High K Strategy Scale had not been adequately confirmed, tested its structure using a national sample from England, and found that a unidimensional model fit more poorly than a model with four correlated first-order factors. The authors also drew attention to several issues in the LH literature that generalized beyond the High K Strategy Scale, including an over-reliance on college samples, a lack of tests for sex differences, small correlations between K and mating effort, lack of validations of psychometric instruments against traditional LH indicators (e.g., pubertal timing or offspring number), and measurement of current environment instead of the early environmental conditions that are the focus of applications of LH theory to human development. Figueredo et al. (2015) responded to this critique, arguing that the psychometrics of human LH strategy were methodologically sound, and Copping and colleagues (2017) stood by their original criticisms in a rejoinder. Readers are encouraged to consult these two papers for more in-depth discussions of the psychometric issues and the degrees to which they have been previously addressed in the literature.

Second, Richardson et al. (2017b) used a bifactor model and data from a large national sample of respondents in the United States to test the structure of LH indicators, as well as whether LH indicators or their underlying dimensions reflected early environmental conditions. The authors' findings suggested that at least two dimensions—mating competition (subsuming, e.g., risk-taking, sensation-seeking, aggression, and number of sexual partners) and Super-K (subsuming, e.g., mental and physical health, conscientiousness, agreeableness, positive attitudes toward children, and social support and pair-bonding)—subsume an array of human LH strategy indicators. Richardson et al. (2017a) replicated these findings in a smaller community sample of young adults. In these studies, childhood and young adult environmental harshness predicted greater mating competition and lesser Super-K

(i.e., a generally faster LH strategy) while early but not late unpredictability predicted lesser Super-K.

Third, Gruijters and Fleuren (2018) pointed out that LH indicators cannot satisfy the assumptions of a reflective measurement model because LH theory is an ultimate level theory and such models assume that factors are the proximate common causes of their indicators. Theories that specify the identities of factors justify the statistical independence among items as well as the independence between items and external criteria (e.g., aspects of early environment). Because ultimate theories do not provide specifications about proximate causation, the authors reasoned, it was unclear how researchers could use LH theory to identify the K factor without conflating ultimate and proximate levels of explanation.

Finally, even before the debate on the psychometrics of LH measures focused on K factors, similar critiques had been leveled at other higher-order constructs. Research in personality had suggested that an overarching personality factor subsuming the Big Five personality dimensions (referred to as the General Factor of Personality, or GFP; Musek, 2007) might reflect an adaptive mechanism to facilitate social behavior (analogous to “g” in intelligence). The assumption is that, as with intelligence, it would never be disadvantageous to have too high a score, and that high GFP individuals would be more reproductively successful. This position, however, has several issues. First, there was evidence to suggest that high GFP is not always advantageous. For instance, those who score highly on Dark Triad measures are often reproductively successful, although they would likely have lower GFP scores (Adams et al., 2014; Carter et al., 2014). Personality traits can also confer differential fitness benefits depending on circumstances (Lowe et al., 2009; Nettle, 2006), and so a linear relationship between GFP and fitness cannot always be assumed. There have also been a number of questions regarding the methodological and statistical basis of GFP (Muncer, 2011) as well as consideration for how such a trait could be selected for (Penke et al., 2007). In short, questions regarding the utility of such all-encompassing higher-order factors were being raised—questions that in many ways were equally applicable to monolithic LH measures that required similar assumptions to GFP.

Evolutionary Genetics Critique

In a more recent critique, Zietsch and Sidari (2020) highlighted several issues in the LH literature drawing largely upon evolutionary genetics, though they also highlighted the poor fit of unidimensional models as discussed above. The authors’ thesis was that adaptive covariation of LH indicators cannot occur except via exceptions to Mendel’s Law of Independent Assortment or developmental plasticity. Zietsch and Sidari went on to argue that none of the exceptions to Mendel’s Law seem promising as sources of adaptive covariance among human traits. Developmental plasticity is also unlikely, they claimed, because studies in statistical genetics have found modest shared environmental effects for most complex traits, and one study did not detect an effect of shared environment on K (Figueroa et al., 2004).

Although full description of, and response to, this most recent critique is outside the scope of the current study, we believe a brief discussion of the authors' arguments will better situate our objectives within the ongoing discussion of the merits of applying LH theory to between-human variance. For a more protracted response that addresses other aspects of Zietsch and Sidari (2020), readers may consult Del Giudice (2020). First, although Zietsch and Sidari identified some important gaps in our understanding of the sources of human LH variation and whether such variation is adaptive, the authors' claims about the promise of applying LH theory to covariance structures in humans appear moot to the extent that proximate causes account for the associations among human LH indicators. That is, past and/or current selection pressures may account for between-species and within-species (human) variance in proximate causes of trait covariance.

In the case of the K-factor, we find the above possibility plausible given that its indicators (e.g., agreeableness, planning and conscientiousness, social support, valuing of children) can be seen as reflecting variations in the default mode network (DMN) that underpins social cognition, including emotion perception, empathy, theory of mind, and morality (Li et al., 2014), as well as remembering the past and thinking about the future (Andrews-Hanna, 2012). Social cognitive variation stemming from the DMN likely manifests as investment in interpersonal relationships with intimate partners and other conspecifics, as well as the planning, empathic responding, moral processing, and responses to social cues required to maintain them. Consistent with this, social connectedness and support, planning, and pair-bonding are among the best indicators of K (Figueredo et al., 2006; Richardson et al., 2017b; Richardson & Sanning, 2017), and research at the interface of evolutionary and personality science has produced evidence consistent with a key role of social cognition in K-factor variance (Van der Linden et al., 2015, 2017, 2018). Providing additional support for this theory, recent research suggests the DMN has undergone recent rapid expansion in humans relative to other primates (Wei et al., 2019; Xu et al., 2019), genes expressed in this human accelerated region appear to differentiate humans from chimpanzees and macaques, and standing variation within humans in these genes is associated with differences in DMN activity, intelligence, mental disorders, such as schizophrenia and autism, and *sociability* (e.g., frequency of friend and family visits; Wei et al., 2019). Taken together, these findings suggest the DMN has undergone recent genetic change, standing variation in DMN genes may be under selection today, and some K-factor indicators reflect variation in DMN genes.

Zietsch and Sidari also concluded that there is little promise in applications of LH theory to plasticity in humans because “in twin studies the shared environment has not been consistently estimated to account for a substantial proportion of variation in any of the aforementioned traits” (2020:531). We find that the authors' conviction is much too strong and inconsistent with recent research examining traits often used as LH indicators, some of which the first author conducted himself. For instance, Zietsch et al. (2010) found that genetic and shared environmental factors explained 33% and 29%, respectively, of the variance in risky sexual behavior in a large sample of young adult twins. Moreover, Verweij et al. (2009) similarly found that shared environmental factors accounted for 33%

and 12% of the variance in retrospectively measured adolescent risky sexual behavior and misconduct, respectively. Other studies have found modest shared environmental effects on age at sexual debut (Bricker et al., 2006; Dunne et al., 1997; Harden et al., 2008; Mustanski et al., 2007; Waldron et al., 2007) and fertility (Kohler et al., 1999; but see Rodgers et al., 2001), and a definitive meta-analysis of virtually all twin and family studies (Polderman et al., 2015) reported modest shared environmental effects on antisocial behavior (and see Maes et al., 2007), education, family relationships, social relationships, and temperament and personality. Shared environmental effects on prosociality have also been documented, and most of the association between parental positivity and prosocial behavior was accounted for by shared environmental factors (Knafo & Plomin, 2006). These studies indicate that many variations in human LH do reflect the environments shared by siblings, consistent with a role of early experiences in LH strategy development.

Notably, the authors' skepticism of theories of LH plasticity is partly based on a single study that estimated the shared environmental variance in K using ACE modeling (Figueredo et al., 2004), a popular statistical genetic methodology for decomposing the phenotypic variance into additive genetic (A), shared environmental (C), and nonshared environmental (E) components (see Neale & Cardon, 2013). Thus it is worth noting that there are empirical and theoretical reasons for caution in interpreting ACE estimates based only on second-order K-factors. First, human LH appears multidimensional, and K is only one potential source of variance in LH indicators (as previously discussed). Second, assuming a genetic watershed, the factor structure of genetic influences on phenotypic indicators may have fewer dimensions, and the associations among them may be larger, than the factor structure of environmental influences. If true, this would imply that common factors (and scale scores) will have higher estimates of heritability than their indicators (i.e., items or subscale scores). That is, because K-factor indicators may share fewer genetic factors than environmental factors, and reflect them to a greater extent, second-order K-factors will tend to largely reflect shared genetic factors while environmental sources of variance will tend to be captured by the first-order residuals. This would mean that researchers searching for environmental influences on LH indicators should examine first-order factors or perhaps use bifactor models to examine first-order residuals in addition to second-order factors (see Chen et al., 2006).

Recognizing the potential importance of examining specific domains in addition to K, Figueredo and colleagues previously specified and tested hybrid models that include the K-factor as well as a developmental cascade among its indicators that reflected environmental influences (e.g., Garcia et al., 2016). Importantly, the appropriateness of interpreting ACE estimates based on second-order K-factors can be addressed empirically via (a) explorations of the genetic and environmental structure of K-factor indicators, which can allow researchers to determine whether genetic sources of covariance among LH indicators are lower-dimensional than environmental sources and have larger effects, and (b) comparisons of common and independent pathway ACE models, which can allow researchers to determine whether a K mediates genetic and environmental

influences on its indicators (Franić et al., 2013). If K does explain both types of influences on indicators, then information about environmental influences is not lost via a more exclusive focus on second-order K-factors.

Next Steps

As noted by Richardson et al. (2017b), determining dimensionality and degree of measurement invariance are key first steps to establishing the validity of a latent construct. If K is indeed a proximate cause of such LH indicators as somatic, parental, and community integrative effort, as theorized above, it should impact the same indicators to the same extent across populations, groups, and settings. Importantly, Copping and Richardson (2020) recently reviewed theory and evidence (discussed further below) suggesting sex differences should be a primary focus in efforts to evaluate the validity of constructs used in psychometric applications of LH theory to human variability. If the structure of K-factor indicators is invariant between the sexes, this increases the plausibility of the proximate cause assumption about K and warrants more in-depth evaluation of its validity.

Sex Differences in Human LH Indicators

Copping and Richardson (2020) reviewed contemporary LH literature empirically examining sex differences. Sex differences are commonly studied in evolutionary psychology as theory recognizes the importance of mechanisms such as sexual (Darwin, 1871) and social (West-Eberhard, 1979) selection, which focus on adaptations pertaining to mating competition and resource acquisition, respectively. Variation between the sexes is also linked to differential levels of parental investment in offspring that reflect costs and benefits of strategies linked to seeking/acquiring mates and rearing offspring (Trivers, 1972). Some scholars also suggest that sex differences can evolve in ways that are not simply consequences of male-male competition and advocate for broader definitions of sexual selection (often referred to as sex-dependent selection; Campbell, 2009; Carranza, 2009). Female fitness could be increased through selection on mechanisms that can increase female survival, and thus the survival of their offspring, given the substantive burden of female parental investment (Sear & Mace, 2008). Males may also be more variable across many biological and psychological traits (referred to as the Greater Male Variability Hypothesis), potentially as a result of differential parental investment strategies (Hill, 2017). Putting the specifics of mechanisms aside, scholars accept that sex differences emerge from pressures surrounding mating and survival, two key features that underpin most trade-offs central to LH theory (Stearns, 1992).

Given the above, sex differences can be expected to manifest in most key first-order LH domains, such as measures of childhood attachment, cognitions, reproductive behaviors, and mating/competition related behaviors. Consequently, the same can be expected of second-order factors (constructs such as the K factor and GFP),

which subsume these lower-order domains. Although variation in first- and second-order LH dimensions can result from varying environmental circumstances (harshness and unpredictability; Ellis et al., 2009), these are likely to impact on men and women simultaneously. Thus the magnitude of sex differences within LH domains should largely remain constant. It is beyond the scope of this paper to summarize the literary body substantiating these claims and interested readers should consult Copping and Richardson (2020) for a narrative review of the contemporary evidence base pertinent to them.

Despite the fact that biological sex is a key source of variation within LH-related constructs, there is a surprising paucity to its consideration within theoretical and empirical LH models (for examples of exceptions, see Del Giudice, 2009; James et al., 2012). Where sex is considered or discussed, it does not always manifest in empirical tests of theory. This therefore raises the question about how exactly sex should be treated in LH-informed research designs. Copping and Richardson (2020) made several suggestions as to how we might proceed with accounting for sex.

First, measurement invariance must be considered. It is common for researchers to use measures of latent constructs and assume that emerging differences between the sexes (or any group) in measures of central tendency must be meaningful. However, unless there is empirical evidence that the underlying factor structure, loadings, and item intercepts are constant across groups, we risk interpreting mean differences that are attributable to item biases or irrelevant test-related errors (i.e., the differences are not actually meaningful properties of group membership; Steinmetz, 2013). Invariance testing is an often-overlooked step in the construct validation process.

Although there are several levels of invariance testing, the three that are of import here are configural (uniformity of the construct structure), metric (uniformity of loadings), and scalar (uniformity of intercepts). Achievement of configural invariance is an absolute necessity for any meaningful comparison. Historical practice focused on achieving at least metric invariance (often partially) for any group comparison. This however may not be enough and can result in errors of prediction, particularly in data structures of increasing complexity (Hsiao & Lai, 2018). Indeed, lack of scalar invariance can also have a large impact in predictive models (Steinmetz, 2013). Consequently, in LH models, there is a need for stricter analysis of invariance if we are to examine differences between groups seriously. Multigroup Confirmatory Factor Analysis (MGCFA), however, can be used effectively if there is at least partial scalar invariance (since mean estimates can be adjusted by allowing free variation of non-invariant intercepts), so effective use of this methodology can be of great assistance to researchers (see Richardson et al., 2017b; Wang et al., 2018). Failure to establish invariance may lead us to erroneous conclusions where sex is concerned. In an LH context, it may suggest that one sex is “faster/slower” than the other even though the metric for each sex is not directly comparable.

An alternative solution, particularly where measurement invariance cannot be established, may be to treat the sexes separately for the purposes of modeling. This approach can be illuminating. Sex differences in means and variances don't always complete the story. Often what may be missing is an analysis of underlying differences in patterns of relationships between key LH constructs. There is good reason

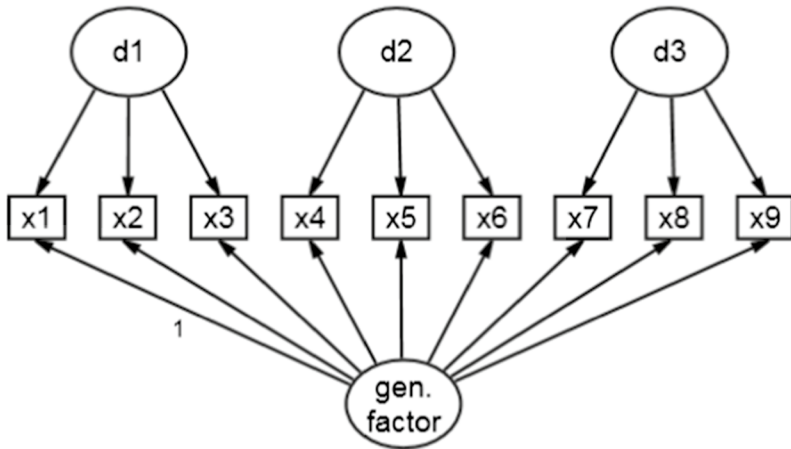
to suggest that LH strategies over the life course differ between the sexes (see James et al., 2012 for an example), with outcomes often linked to different proximal or distal causes at varying magnitudes. These subtle, but informative, effects would be missed if we simply examined aggregated sample data. Such models would illuminate which pressures, and the outcomes they impact, are important to men and women, even if the means and variances within constructs used to construct the model are not directly comparable. Such analyses may lead us to more informative conclusions whose implementation could have major theoretical and practical interventionist implications.

Current Study

In this study, we reanalyze the Richardson et al. (2017b) data using exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009) to determine whether the structure of LH indicators is invariant by sex. ESEM is an approach that is useful when researchers do not wish to specify the number of factors a priori or assume there are no cross-loadings (i.e., unlike in confirmatory factor analysis). ESEM differs from traditional exploratory factor analysis (EFA) in that (a) it produces fit statistics that can be used to evaluate model correspondence to data, meaning the number of factors can be determined partly on the basis of model fit; (b) standard errors for the rotated solution as well as residual covariances are estimable; and (c) parameter constraints can be imposed on the model to test for measurement invariance.

A challenging finding reported in Richardson, Sanning, et al. is that mating competition was statistically independent of Super-K. This is not consistent with most applications of LH theory to humans, which predict a trade-off at the phenotypic level (i.e., an inverse correlation; e.g., see Kruger, 2017) between dimensions subsuming mating effort and somatic and parental effort. One possibility unexamined by the authors is that the independence between mating competition and Super-K was an artifact of the orthogonality constraints imposed on bifactor models (discussed further in “Methods”). ESEM is appropriate for the current study because in addition to allowing us to test for invariance by sex, it enables us to examine whether a LH indicator structure similar to that found in Richardson et al. (2017b) emerges in the absence of theory-driven model constraints and fits the data as well as the authors’ confirmatory bifactor model. In addition to providing additional insight into the structure of human LH indicators, the current study serves as a case example in the general use of ESEM as well as an exploration of the convergence between ESEM and bifactor model results. It may, therefore, provide some guidance to researchers wishing to evaluate the sensitivity of their results to constraints placed on confirmatory models.

Bifactor Model



Second-order Model

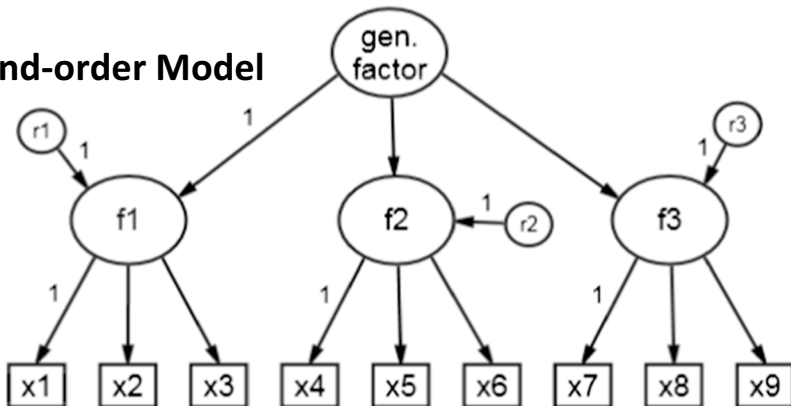


Fig. 1 Bifactor versus second-order factor model. *d*=domain specific; *f*=first-order factor; *gen* general, *r* residual (Richardson et al., 2017: Fig. 1)

Methods

Revisiting the Bifactor Model

As described in Richardson et al. (2017b), most factor-analytic studies of human LH strategy have relied on the second-order factor model (Fig. 1), which includes higher-order or “general” constructs (e.g., K or Super-K factors) that subsume multiple first-order constructs (e.g., warmth of relationships with parents). These models typically specify the variances in first-order constructs unexplained by higher-order factors (i.e., the first-order factor residuals) as orthogonal to second-order factors and other first-order factors. Moreover, the first-order factor residuals are usually ignored (Chen et al., 2006). But researchers are sometimes interested in general

factors and also factors that explain unique variance in domains over and above that explained by general factors. They may want to predict outcomes such as academic achievement with general factors (e.g., intelligence) and also unique variances in one or more specific domains (e.g., verbal ability). Researchers may also wish to estimate the effects of general as well as domain-specific factors on their items or determine if domain-specific variances are significant after accounting for a general factor. As Chen et al. (2006) noted, the absence of domain-specific variance in second-order models often goes unnoticed. In sum, bifactor models (see Fig. 1) allow for estimation of effects of general factors (e.g., Super-K) on their indicators and also examination of whether, and to what extent, variance in indicators that is not explained by general factors reflects other constructs or dimensions.

Although the pragmatic use of bifactor models for partitioning variance into general and specific factors (e.g., to determine if specific factors predict criteria of interest holding the general factor constant; Chen et al., 2006) has generated little controversy, greater scrutiny has been leveled at efforts to determine whether bifactor or second-order models better account for the associations among items or first-order constructs (e.g., domains of psychopathology; Bonifay et al., 2017; Gignac & Kretzschmar, 2017; Mansolf & Reise, 2017; Snyder et al., 2017). First, bifactor models tend to fit better than second-order models when the true model has error covariances and/or cross-loadings (Mansolf & Reise, 2017). This might lead researchers to incorrectly select bifactor models over second-order models. Second, the general and specific factors in bifactor models do not always seem theoretically plausible—can we attribute unique identities to the specific factors?

In our view, the first issue is less likely applicable to Richardson et al. (2017b)'s findings because they did not compare bifactor and second-order models, but instead used the bifactor model to examine variance in Super-K indicators not accounted for by the general factor. The authors did claim that a two-dimensional structure fit the Super-K and mating competition data better than a single dimension. It seems unlikely (to us) that specifying a second-order or other model of the Super-K indicators would change the association between mating competition and Super-K dramatically enough to make a single dimension plausible (i.e., from consistent with the null to, say, $r=0.80$); however, additional research is needed to determine the extent to which the structure reported in Richardson et al. (2017b) might be an artifact of the orthogonality constraints imposed on their bifactor model.

The second issue more clearly applies to Richardson et al. (2017b) because, as mentioned above, the authors examined the variance in Super-K indicators not explained by the general factor (i.e., the variance captured by the specific factors). As noted by Bonifay et al. (2017), specific factors may not have identities that are plausibly distinct from general factors. This implies that in Richardson et al. (2017b), the orthogonality constraints imposed between Super-K and the specific factors could be questionable. It is also possible that these orthogonality constraints forced some Super-K indicator variance, which would have been captured by the general factor in second-order models, into the specific factors. If this variance was shared with mating competition, perhaps the constraints on the bifactor model reduced the estimated magnitude of the covariance between mating competition and Super-K.

In the current study, we use ESEM because it allows us to test for invariance by sex while simultaneously addressing this possibility.

Data

This study reanalyzed publicly available national data from the surveys of Midlife in the United States (MIDUS; $n=4,244$). The MIDUS surveys investigated the role of behavioral, psychological, and social factors in accounting for age-related variations in health and well-being. We used the data from 2004–2006 (Round 2) and sample demographics are described in Richardson et al. (2017b). For more information about the MIDUS samples, please visit <http://www.midus.wisc.edu/midus1/index.php>.

Instruments

This study examined factor and index scores produced by Richardson et al. (2017b) using data associated with 19 Round 2 indicators¹ (measured with 73 items) of middle adult LH strategy. A full list of the items and scale scores used is presented in the ESM (reproduced from Richardson et al., 2017b), along with the associated MIDUS labels, constructs measured, and item contents. The original report also contains citations to reviews of LH measures or other publications where the use of each indicator is substantiated. For further information about the MIDUS scales, including coefficient α for the MIDUS samples and references to validation studies, please access the MIDUS I and II scales information at <http://www.midus.wisc.edu/midus1/index.php> and <http://www.midus.wisc.edu/midus2/project1/>.

Analysis

We conducted our ESEM analyses (Asparouhov & Muthén, 2009) using the MPlus 8 software package and the robust weighted least squares (WLSMV) estimator because two items were ordinal (Muthén et al., 1997). We conducted all significance tests at the 0.005 level to protect against Type I errors (Benjamin et al., 2018) and used $\beta > 0.30$ as the threshold for interpreting factor loadings (Kline, 2015). We used the delta parameterization and oblique rotations (Geomin). We used multiple group ESEM for our invariance testing. For descriptions of how ESEM and multiple group ESEM models are parameterized, see the Mplus User Guide Version 8, examples 4.1 and 5.27 (http://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf).

Model fit. We used a variety of fit indices because they provide different information about model fit. We considered the substantive meaningfulness of the model, nonsignificance of the χ^2 likelihood ratio statistic (Bollen, 1989), Tucker–Lewis

¹ There are 21 scores total because two indicators selected in the original study are two-dimensional.

Table 1 ESEM model fit information

Model	$\chi^2(df)$	CFI	TLI	RMSEA(95% CI)
2 Factors	7984.900 (187)*	.747	.716	.099 (.097, .101)
3 Factors	5285.624 (171)*	.834	.796	.084 (.082, .086)
4 Factors	3151.314 (156)*	.903	.869	.067 (.065, .069)
5 Factors	2257.646 (142)*	.931	.898	.059 (.057, .061)
4 Factors b	1455.717 (146)*	.957	.939	.046 (.044, .050)

*Statistically significant at $p < .005$

index (TLI) and comparative fit index (CFI) greater than 0.95 (Byrne, 2001; Hu & Bentler, 1999), and root means square error of approximation (RMSEA) values of less than 0.05 (Browne & Cudeck, 1993) as evidence of acceptable fit to the data. In this study, we used the corrected χ^2 difference test ($\Delta\chi^2$) and differences in CFI (Δ CFI) that exceeded 0.01 (Cheung & Rensvold, 2002) to evaluate whether invariance held across the sexes. We used Δ CFI partly because we expected χ^2 difference tests to be significant even in cases when differences in fit were trivial given the large size of our sample.

Hypothesized model. Our modeling approach consisted of two interrelated steps. First, we used ESEM to test the structure of the LH indicators in the sample as a whole. We began by specifying a single-factor measurement model and then compared it with models with increasing numbers of factors, ceasing when model fit was observed as excellent or solutions did not make substantive sense. This allowed us to examine whether a LH indicator structure similar to that found in Richardson et al. (2017b) emerged in the absence of theory-driven model constraints. Our approach was exploratory and did not specify the number of factors that would subsume our LH indicators. However, we did specify a substance use/abuse factor in light of past research that reported a unidimensional structure of substance use/abuse indicators (Richardson et al., 2014, 2016, 2017a, 2017b). Thus, the model included an ESEM LH portion as well as a confirmatory substance use/abuse portion. We did not make any a priori predictions about the presence or absence of cross-loadings in the ESEM. In our second step, we used multiple group ESEM to test for configural, metric, scalar invariance across the sexes. We also tested for sex differences in factor means and variances.

Results

ESEM

This study used ESEM to reexamine nationally representative middle adult data analyzed in Richardson et al. (2017b). Our objectives were to (a) use an exploratory approach to determine whether a structure similar to that reported in the original study emerged, as well as (b) test for measurement invariance by sex. Regarding (a),

Table 2 Standardized factor loadings from final model (4b)

	Super-K	Mating competition	GFP	Substance use/abuse
Extroversion	.14		.74	
Openness	–.14		.95	
Conscientiousness	.15	–.04	.53	
Agreeableness	.01	–.52	.45	
Neuroticism	–.49			
Pair-bond	.52		–.11	
Positive affect	.57		.12	
Social support	.59	–.26	.14	
Education	.47			
Mental/physical health	.80	.30	–.01	
Aggression	–.34	.44		
No. of Sexual partners		.30	.15	
Risk-taking		.41	.19	
Neighborhood quality	.23	.44	.08	
Alcohol abuse	.14	.36		
Alcohol use				.33
Smoking				.29
Illicit drug use				.46
Illicit drug abuse				.37
Sexuality	.19	.28	.18	
No. of Children	.07	–.20	–.07	

Table displays standardized loadings associated with statistically significant unstandardized effects. Standardized effects larger than .30 are bolded

we sought to determine whether statistical independence between mating competition and Super-K could have been an artifact of the a priori constraints imposed on bifactor models. To achieve our objectives, we conducted an ESEM analysis starting with a single LH factor as well as the substance use/abuse factor. We tested this model, and it fit the data poorly (e.g., CFI=0.747; see Table 1 for complete fit information). We then increased the number of factors subsuming the LH indicators to two (three factors in total; CFI=0.834), three (four factors in total; CFI=0.903), and four (five factors in total; CFI=0.931). The three latent LH variables in the four-factor model were readily interpretable as Super-K, mating competition, and the general factor of personality (GFP); the fourth factor was liability to substance use/abuse (Table 2). The same indicators loaded substantially on these factors as in the original study (Richardson et al., 2017b). For instance, social support, mental/physical health, education, pair-bonding, and neuroticism loaded substantially on Super-K.

The five-factor model was difficult to interpret. For one of the LH factors, only one indicator (agreeableness) had a large loading ($\beta=0.67$), and the next highest

Table 3 Factor intercorrelations

	Super-K	Mating compe- tition	GFP
Super-K	1		
Mating competition	-.20	1	
GFP	.57	.02 ^{ns}	1
Substance Use/abuse	-.48	.36	-.05 ^{ns}

loading indicator was neuroticism ($\beta=0.36$). Given the sizes and pattern of the loadings, the identity of this factor was unclear (e.g., what causes higher agreeableness *and* neuroticism?). We speculate that one possibility, given the agreeableness loading is nearly twice the size of the next largest loading, is that this factor is due to effects of a construct closely related to agreeableness on other indicators. Perhaps the small loading of aggression ($\beta = -.21$) on this factor is consistent with this interpretation. Ultimately, we lacked a clear theoretical identity for this factor and therefore returned to the four-factor model to observe modification indices (MIs) for sources of strain. We found relatively large MIs (> 100) associated with five residual covariances, which were specified on the basis of modification indices and substantive theory in the original study (sexuality \leftrightarrow # sexual partners; sexuality \leftrightarrow pair-bonding; neighborhood quality \leftrightarrow smoking; neighborhood quality \leftrightarrow aggression; agreeableness \leftrightarrow neuroticism), and added these parameters. We also identified two additional MIs that were relatively large (> 100) and substantively plausible (pair-bonding \leftrightarrow # sexual partners; pair-bonding \leftrightarrow illicit drug use). We added these parameters to the model (4 factors b) and tested it. Fit was acceptable and very similar to that observed for the final bifactor model in the original study ($\Delta CFI=0.002$). Given that fit was adequate and no relatively large MIs remained, we accepted this as our final model of the associations among the LH indicators.

Next, we observed and interpreted the correlations among the factors (see Table 3) to determine the plausibility of a single dimension of human LH strategy, whether Super-K and mating competition were inversely correlated, and whether substance use/abuse was associated with greater mating competition and lesser Super-K, as in the original study. We observed a large positive correlation between Super-K and GFP ($r=0.57$), but a single LH dimension was implausible because mating competition and GFP were uncorrelated and because we observed a small inverse association between Super-K and mating competition ($r = -0.20$). We found moderate positive and negative associations between liability to substance use/abuse

Table 4 Invariance testing model fit information

Model	χ^2 (df)	χ^2_{males}	χ^2_{females}	CFI	TLI	RMSEA (95% CI)
Scalar model 1	1886.911 (364)*	981.972	904.939	.950	.943	.044 (.042, .046)
Scalar model 2	1607.040 (351)*	858.007	749.033	.959	.951	.041 (.039, .043)
Variance model	1349.786 (358)*	716.033	633.753	.968	.962	.036 (.034, .038)

*Statistically significant at $p < .005$

Table 5 Female means and variances

Latent Variables	σ^2	99.5% CI	\bar{x}	99.5% CI
Super-K	1.046*	0.913, 1.178	-0.054	-0.173, 0.066
GFP	1.047*	0.899, 1.195	-0.146*	-0.235, -0.057
Mating Competition	0.864*	0.710, 1.018	-1.243*	-1.403, -1.082
Substance use/abuse	0.017*	0.010, 0.023	0.014	-0.014, 0.043

CI Confidence interval. Female means are estimated as differences from the male means, which are fixed to zero to identify the model

* $p < .005$

and mating competition ($r=0.36$) and Super-K ($r=-0.48$), respectively. Notably, alcohol abuse only reflected mating competition ($\beta=0.36$) and neuroticism only reflected Super-K (not GFP; $\beta=-0.49$).

Measurement Invariance

Next we tested our final exploratory model (4 factors b) for measurement invariance by sex. Loadings, thresholds, and intercepts/thresholds were constrained equal across the sexes by default. We found that this model (scalar model 1) fit the data marginally well (e.g., CFI=0.950; see Table 4). Several modification indices were relatively large (>30), suggesting the intercepts associated with neuroticism, covitality, pair-bonding, and neighborhood quality, as well as the thresholds associated with number of children, were not invariant between the sexes. We freed these intercepts and thresholds, tested the resulting model, and found that fit improved significantly ($\Delta\chi^2=293.669[13]$, $p < 0.001$; Δ CFI=0.009).

Next, we sought to test for sex differences in the latent variable means and variances. Mean liability to substance use ($p=0.192$) and Super-K ($p=0.246$) did not significantly differ between males and females. In contrast, both mean GFP and mating competition were significantly lower among females (p values <0.001). To test the equality of ESEM factor variances in the context of multiple groups, factor variances and covariances must both be constrained equal. Recall that the liability to substance use/abuse factor was not an ESEM factor. We were able, therefore, to constrain its variance equal across the sexes without modifying its covariances with other factors. We constrained the Super-K, GFP, and mating competition variances and covariances, as well as the liability to substance use variance, equal across the sexes (variance model). These constraints did not significantly degrade model fit ($\Delta\chi^2=19.437[7]$, $p=0.006$; Δ CFI=0.001), suggesting that none of the factor variances significantly differed between males and females. Female means and variances are displayed in Table 5. Male parameters are not displayed because males served as the reference group in this study and their means and variances were fixed to zero and one, respectively, to identify the model. Females served as the non-reference groups and their means and variances were freely estimated. Tests of the null hypothesis that the female means are zero provided comparison of male and female means. As noted above, parameter constraints were used to compare the male and female variances.

Discussion

This study used an exploratory approach to reexamine nationally representative middle adult data analyzed in Richardson et al. (2017b). ESEM findings in this study were similar to those derived from the bifactor model in the original study. On the basis of model fit and substantive rationale, we accepted a four-factor solution as the best reproducer of the associations among the LH indicators. The ESEM factors were Super-K, mating competition, and GFP; the fourth (confirmatory) factor was liability to substance use/abuse. No covitality (i.e., health and mental health) factor emerged, but this was unsurprising given that only two indicator loadings on this factor (neuroticism and health/mental health) were greater than $\beta=0.30$ in the original study.

We found an inverse association between mating competition and Super-K, suggesting the statistical independence between these factors in the original study may have been an artifact of constraints on the bifactor model. This inverse association is also consistent with a trade-off between mating effort and somatic, parental, and community integrative effort. We also observed a large positive correlation between Super-K and GFP, consistent with earlier research specifying the latter as an indicator of the former (for a review, see Figueredo et al., 2006). We again found that neuroticism only loaded on Super-K, suggesting the GFP may not subsume this construct. This result accords well with the criticisms of GFP discussed above in our review (e.g., Muncer, 2011). Finally, liability to substance use/abuse reflected greater mating competition and lesser Super-K, and this is also consistent with prior research (e.g., Richardson et al., 2016).

The results produced by our alternative approach suggest the finding of multidimensionality in Richardson et al. (2017b) is robust to the modeling approach employed—the evidence against a single dimension does not seem to be an artifact of constraints on the bifactor model. In particular, here we found that GFP was uncorrelated with mating competition, and a small inverse correlation linked mating competition and Super-K. This evidence is consistent with patterns of correlations observed in other studies that examined broad selections of LH indicators (e.g., Kogan et al., 2014; Međedović, 2018) and indicates that researchers should cease interpreting single scores (e.g., Mini-K and other K-factor scores) as empirical proxies for LH speed. In psychometrics, construct validity hinges on internal structure and the nomological net (for further discussion, see Richardson et al., 2017b). The first problem with interpreting K-factor scores as LH speed is that, although K-factor scores have been linked to other constructs in a way that is consistent with LH theory as applied to humans, they do not provide sufficient information about such LH indicators as pubertal timing, age at sexual debut, mating effort, risky proneness, and antisocial behavior to serve as their empirical proxy. Again, to be clear, K-factor scores have been *linked* to these variables. However, K-factors are not useful empirical proxies for these variables because they don't satisfy the assumption of local independence to an acceptable degree. Because of this, K-factor scores cannot be interpreted as measuring LH speed.

The second problem with interpreting K-factor scores as LH speed is that, as discussed previously, LH speed has never been theorized as a proximate cause of anything in biology. It is a pattern of covariation among LH indicators that is explained at the ultimate level (see Gruijters & Fleuren, 2018). Thus, there simply aren't adequate psychometric or theoretical grounds for interpreting K-factor scores as LH speed. We realize the transition away from interpreting K-factor scores in this way may be onerous for some, particularly if they used this approach in the past. Indeed, a large literature has now interpreted K-factor scores in this way (see Figueredo et al., 2014). Given this, we caution researchers to avoid arguing from the size of the literature or from "previous research used this approach." Neither of these arguments is compelling because they do not remediate any of the problems identified above.

The current study also tested for sex differences in the structure of LH indicators, and several findings are notable. First, we found that configural and metric invariance held for the LH dimensions. These findings suggest latent LH variables may have equivalent effects in each sex and are consistent with our theory that K is a proximate cause of LH indicator covariance that emerges from DMN variations. Second, we found that the intercepts for covitality, neuroticism, pair-bonding, and neighborhood quality, as well as the thresholds for number of children, varied between the sexes. These findings imply that mean sex differences in these LH indicators could not be fully attributed to LH factors and are consistent with the Richardson et al. (2017c) finding that the Mini-K was partially scalar invariant by sex. If we had computed a sum or mean score across the Super-K indicators and tested for a mean sex differences, ignoring these noninvariant intercepts, this would have likely biased the mean difference and could have led us to incorrectly conclude that there were mean sex differences on Super-K. Indeed, we found no evidence of sex differences in the Super-K means or variances. GFP was significantly lower in females, but its variance did not differ significantly by sex. Mating competition was significantly greater on average in males relative to females, but its variance also did not differ significantly by sex. Taken together, these findings cohere well with our suggestion that measurement invariance testing is crucial in efforts to evaluate the validity of LH constructs, as well as in efforts to detect sex differences (Copping & Richardson, 2020; Wang et al., 2018).

Limitations

The limitations associated with this study mirror those in the original (Richardson et al., 2017b). In particular, this study is limited by the use of self-report data and assumed, at least in part, the validity of the many psychological constructs used as indicators of LH strategy. These constructs have been studied extensively with factor modeling, and the original study established the unidimensionality of each scale before saving factor scores. The original study also used a method (i.e., Bartlett's) that produces unbiased estimates of the true factor scores (Hershberger, 2005), and most of our LH domains contained one or more survey-developer-created scale scores, which were produced using validated scales. However, the validities of some

constructs may deserve more scrutiny. For instance, testing all our scales for invariance by sex was infeasible, and thus we could not be sure that some of scales used to measure LH indicators did not exhibit some degree of noninvariance by sex. This could have led us to detect some noninvariance that was due to measurement bias and not true sex differences in structure. Importantly, however, our conclusions regarding the latent means and variances are not imperiled by this limitation—we left noninvariant intercepts free to vary between the sexes, and as a consequence, they did not impact latent level parameters (see Wang et al., 2018).

Future Directions

Our use of ESEM demonstrates the methodological strength in additional exploratory work to evaluate the unidimensionality of constructs. While this study retains its focus on Super-K style constructs, there is no reason as to why this approach cannot be widely employed when assessing other latent psychological constructs (e.g., sociosexual orientation) relevant to evolutionary approaches to human behavior. Researchers should aim to provide evidence that unidimensional solutions are quantifiably better than multidimensional alternatives by testing between alternative (but theoretically substantive) model specifications and quantifying relationships between relevant subordinate factors/traits. This would help establish a firm validity case for measurement instruments. Without such evidence, unidimensional measures should be interpreted cautiously.

As discussed above, the findings presented here suggest that K-factor scores are not adequate empirical proxies for LH speed. They also suggest that scores likely introduce bias when they are used to compare the sexes. To minimize bias, future research should identify indicators of latent LH variables that are most consistent with theory (e.g., in the case of Super-K, these are likely indicators somatic, parental, and community integrative effort) and invariant across the sexes and other groups before scores are used. Researchers should consider specifying indicators with loose theoretical ties to LH factor identities and/or small or noninvariant loadings as correlates of LH factors rather than their indicators.

Future research may also use generalized network psychometrics (see Epskamp et al., 2017) to explore residual structures connecting LH indicators while simultaneously modeling latent LH dimensions. This approach would be consistent with the residual covariance detected in the current study as well as theoretical and empirical work previously carried out by Figueredo and colleagues (Garcia et al., 2016). Notably, an appreciable degree of invariance held for the LH dimensions, and no Super-K or mating competition loadings appeared to vary by sex. This evidence suggests LH dimensions may have equivalent effects in each sex and is consistent with our theory that K may be a proximate cause of LH indicator covariance that emerges from DMN variations. Further efforts to test this theory are warranted.

As discussed previously, future studies can further evaluate our proximate cause theory of the K-factor by testing whether genetic and environmental factors influence K-factor indicators largely through K. The logic here is the same as with multiple indicators and multiple causes (MIMIC) models (Jöreskog & Goldberger, 1975).² A proximate cause should mediate effects of external criterion variables on its indicators. Whereas the typical MIMIC model includes upstream sources of indicator variance that are phenotypic, in statistical genetic comparisons of common and independent pathways models the upstream sources of indicator variance are unmeasured and decomposed into the ACE components using parameter constraints (see Neale & Cardon, 2013). In both cases, we would expect few direct effects (i.e., effects unmediated by K) on K-factor indicators if K was indeed their proximate cause. If this is correct and the common pathways model fits as well as the independent pathways model, the Zietch and Sidari (2020) argument that plasticity plays little role in individual differences in K-factor indicators might be correct given the lack of a shared environmental effect in Figueredo et al. (2004). If instead the independent pathways model fits best, suggesting some effects of genetic and environmental factors are not mediated by K, this would be consistent with the Garcia et al. (2016) model that included effects on K-factor indicators that were not accounted for by K. To the degree that such effects exist, K-factor scores may be less useful and studies of this LH dimension may need to proceed via latent variable modeling. Importantly, researchers can evaluate the practical significance of partial invariance (scalar in this case), or the degree to which ignoring noninvariance and computing scores introduces bias (for a tutorial and R script for doing so, see Lai et al., 2019). Moreover, as the number and size of direct effects (i.e., effects unmediated by K) on K-factor indicators increases, the realist assumption that K is their proximate cause becomes less tenable. Taken together, the current findings of scalar noninvariance suggest that researchers should select indicators that best reflect K theoretically, demonstrate measurement invariance across the sexes and other groups, and then test MIMIC models or compare common and independent pathways models.

Several other directions for future research are worth highlighting. First, many studies of the K-factor have been conducted using university samples and thus more studies of diverse samples are still needed. Second, a perhaps disproportionate number of the extant psychometric studies of human LH emerged from one lab (that of Figueredo and colleagues), and thus contributions from more independent groups would likely strengthen the literature. Third, while Figueredo and colleagues took pains to address some potential threats to validity (e.g., social desirability bias;

² Measurement invariance testing via multiple groups structural equation modeling (SEM) is closely related to MIMIC modeling. If strict scalar invariance by sex holds for the K-factor in the former approach (i.e., all loadings and intercepts invariant), then there will be no direct effects of sex on K-factor indicators in the latter approach, as well as no moderation of K-factor loadings by sex. In a MIMIC model containing the K-factor, direct effects on reflective indicators of K would represent evidence of scalar noninvariance, or that intercepts vary by sex. By entering a $K \times \text{sex}$ interaction term into the MIMIC model, researchers can also test whether sex appears to moderate effects of K on its indicators. Moderation by sex, in this case, is evidence of metric noninvariance or that loadings vary between the sexes. Multiple groups SEM offers several advantages beyond MIMIC models, including the possibility of testing for differences in variances between groups and more straightforward testing for differences in loadings and covariances.

Figueredo et al., 2014), other sources of bias, including halo bias and common method bias, have not been adequately addressed. Finally, although more research devoted to detecting evidence consistent with trade-offs is emerging in the literature attendant to both psychosocial and biodemographic indicators (Beall & Schaller, 2019; Međedović, 2019), much more of this kind of work is needed (Black et al., 2017) and studies of the role of ecology in trade-offs would be particularly informative (e.g., see Stiver & Alonzo, 2009). We hope future research can address all these areas in which the human LH literature might be strengthened.

Conclusion

This study used ESEM to examine whether a LH indicator structure similar to that found in Richardson et al. (2017b) emerged in the absence of theory-driven model constraints, as well as to determine whether the structure of LH indicators was invariant by sex. Findings were similar to the original study in that a single LH speed dimension was not plausible, but they suggest that the statistical independence between mating competition and Super-K may have been an artifact of the orthogonality constraints on the bifactor model. In the current study, a small inverse correlation between these variables was detected. Measurement invariance testing revealed evidence of metric invariance, consistent with the theory that K is a proximate cause of its indicators; however, evidence of partial scalar invariance indicates that use of scores likely introduces bias when the sexes are compared. Taken together, findings suggest researchers should cease interpreting K-factor scores as empirical proxies for LH speed, adopt modeling approaches allowing for examination of factor structures concurrent with residual structures connecting LH indicators, and engage in further evaluation of the validity of K via MIMIC and behavioral genetic modeling.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12110-021-09398-5>.

Acknowledgments Publically available data from the MIDUS study was used for this research. Since 1995, the MIDUS study has been funded by the following: John D. and Catherine T. MacArthur Foundation Research Network, National Institute on Aging (P01-AG020166), and National Institute on Aging (U19-AG051426).

References

- Adams, H. A., Luevano, V. X., & Jonason, P. K. (2014). Risky business: Willingness to be caught in an extra-pair relationship, relationship experience and the dark triad. *Personality and Individual Differences*, 66, 204–207. <https://doi.org/10.1016/j.paid.2014.01.008>
- Andrews-Hanna, J. R. (2012). The brain's default network and its adaptive role in internal mentation. *The Neuroscientist*, 18, 251–270. <https://doi.org/10.1177/1073858411403316>
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, 16, 397–438. <https://doi.org/10.1080/10705510903008204>
- Beall, A. T., & Schaller, M. (2019). Evolution, motivation, and the mating/parenting trade-off. *Self and Identity*, 18(1), 39–59. <https://doi.org/10.1080/15298868.2017.1356366>

- Belsky, J., Steinberg, L., & Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child Development*, 62(4), 647–670. <https://doi.org/10.1111/j.1467-8624.1991.tb01558.x>
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., & Cesarini, D. (2018). Redefine statistical significance. *Nature Human Behaviour*, 2(1), 6–25. <https://doi.org/10.1038/s41562-017-0189-z>
- Black, C. J., Figueredo, A. J., & Jacobs, W. J. (2017). Substance, history, and politics: An examination of the conceptual underpinnings of alternative approaches to the life history narrative. *Evolutionary Psychology*, 15(1). <https://doi.org/10.1177/1474704916670402>
- Bollen, K. (1989). *Structural equations with latent variables*. Wiley Series in Probability and Mathematical Statistics. Wiley.
- Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science*, 5(1), 184–186.
- Bricker, J. B., Stallings, M. C., Corley, R. P., Wadsworth, S. J., Bryan, A., Timberlake, D. S., & DeFries, J. C. (2006). Genetic and environmental influences on age at sexual initiation in the Colorado Adoption Project. *Behavior Genetics*, 36(6), 820–832.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.
- Byrne, B. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (1st ed.). Lawrence Erlbaum Associates.
- Campbell, A. (2009). What kind of selection? *Behavioral and Brain Sciences*, 32, 272–273.
- Carranza, J. (2009). Defining sexual selection as sex-dependent selection. *Animal Behaviour*, 77, 749–751.
- Carter, G. L., Campbell, A. C., & Muncer, S. (2014). The dark triad: Beyond a ‘male’ mating strategy. *Personality and Individual Differences*, 56, 159–164.
- Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, 41(2), 189–225.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.
- Copping, L. T., & Richardson, G. B. (2020). Studying sex differences in psychosocial life history indicators. *Evolutionary Psychological Science*, 6(1), 47–59. <https://doi.org/10.1007/s40806-019-00211-2>
- Copping, L. T., Campbell, A., & Muncer, S. (2014). Psychometrics and life history strategy: The structure and validity of the high K strategy scale. *Evolutionary Psychology*, 12(1), 200–222.
- Copping, L. T., Campbell, A., Muncer, S., & Richardson, G. B. (2017). The psychometric evaluation of human life histories: A reply to Figueredo, Cabeza de Baca, Black, Garcia, Fernandes, Wolf, and Woodley (2015). *Evolutionary Psychology*, 15(1). <https://doi.org/10.1177/1474704916663727>
- Darwin, C. (1871). *The descent of man and selection in relation to sex*. Penguin Classics (2004 edition).
- Del Giudice, M. (2009). Sex, attachment and the development of reproductive strategies. *Behavioral and Brain Sciences*, 32, 1–67. <https://doi.org/10.1017/S0140525X09000016>
- Del Giudice, M. (2020). Rethinking the fast-slow continuum of individual differences. *Evolution & Human Behavior*, 41, 536–549. <https://doi.org/10.1016/j.evolhumbehav.2020.05.004>
- Draper, P., & Harpending, H. (1982). Father absence and reproductive strategy: An evolutionary perspective. *Journal of Anthropological Research*, 38(3), 255–273.
- Dunne, M. P., Martin, N. G., Statham, D. J., Slutske, W. S., Dinwiddie, S. H., Bucholz, K. K., & Heath, A. C. (1997). Genetic and environmental contributions to variance in age at first sexual intercourse. *Psychological Science*, 8(3), 211–216.
- Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, 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.
- Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*, 82(4), 904–927.
- Figueredo, A. J., de Baca, T. C., Black, C. J., García, R. A., Fernandes, H. B. F., Wolf, P. S. A., & Anthony, M. (2015). Methodologically sound: Evaluating the psychometric approach to the assessment of human life history [reply to]. *Evolutionary Psychology*, 13(2), 147470491501300200.
- Figueredo, A. J., Vasquez, G., Brumbach, B. H., & Schneider, S. M. (2004). The heritability of life history strategy: The K-factor, covitality, and personality. *Social Biology*, 51(3–4), 121–143.
- Figueredo, A. J., Vásquez, G., Brumbach, B. H., & Schneider, S. M. (2007). The K-factor, covitality, and personality. *Human Nature*, 18(1), 47–73.

- Figueredo, A. J., Vásquez, G., Brumbach, B. H., Schneider, S. M., 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(2), 243–275.
- Figueredo, A. J., Vásquez, G., Brumbach, B. H., Sefcek, J. A., Kirsner, B. R., & Jacobs, W. J. (2005). The K-factor: Individual differences in life history strategy. *Personality and Individual Differences*, 39(8), 1349–1360.
- Figueredo, A. J., Wolf, P. S. A., Olderbak, S. G., Gladden, P. R., Fernandes, H. B. F., Wenner, C., & Jacobs, W. J. (2014). The psychometric assessment of human life history strategy: A meta-analytic construct validation. *Evolutionary Behavioral Sciences*, 8(3), 148–185.
- Franić, S., Dolan, C. V., Borsboom, D., Hudziak, J. J., van Beijsterveldt, C. E., & Boomsma, D. I. (2013). Can genetics help psychometrics? Improving dimensionality assessment through genetic factor modeling. *Psychological Methods*, 18(3), 406–433.
- García, R. A., de Baca, T. C., Black, C. J., Sotomayor-Peterson, M., Smith-Castro, V., & Figueredo, A. J. (2016). Measures of domain-specific resource allocations in life history strategy: Indicators of a latent common factor or ordered developmental sequence? *Journal of Methods and Measurement in the Social Sciences*, 7(1), 23–51.
- Gignac, G. E., & Kretzschmar, A. (2017). Evaluating dimensional distinctness with correlated-factor models: Limitations and suggestions. *Intelligence*, 62, 138–147.
- Grujters, S. L., & Fleuren, B. P. (2018). Measuring the unmeasurable. *Human Nature*, 29, 33–44.
- Harden, K. P., Mendle, J., Hill, J. E., Turkheimer, E., & Emery, R. E. (2008). Rethinking timing of first sex and delinquency. *Journal of Youth and Adolescence*, 37(4), 373–385. <https://doi.org/10.1007/s10964-007-9228-9>
- Hershberger, S. L. (2005). Factor scores. In B. S. Everitt (Ed.), *Encyclopedia of statistics in behavioral science* (pp. 636–644). Wiley.
- Hill, T. P. (2017). An elementary theory for the variability hypothesis. *New York Journal of Mathematics*, 23, 1641–1655.
- Hsiao, Y. Y., & Lai, M. H. C. (2018). The impact of partial measurement invariance on testing moderation for single and multi-level data. *Frontiers in Psychology*, 9, 740. <https://doi.org/10.3389/fpsyg.2018.00740>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- James, J., Ellis, B. J., Schlomer, G. L., & Garber, J. (2012). Sex specific pathways to early puberty, sexual debut and sexual risk taking: Test of an integrated evolutionary-developmental model. *Developmental Psychology*, 48, 687–702. <https://doi.org/10.1037/a0026427>
- Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, 70(351a), 631–639.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford.
- Kogan, S. M., Cho, J., Simons, L. G., Allen, K. A., Beach, S. R. H., Simons, R. L., & Gibbons, F. X. (2014). Pubertal timing and sexual risk behaviors among rural African American male youth: Testing a model based on life history theory. *Archives of Sexual Behavior*, 44(3), 609–618. <https://doi.org/10.1007/s10508-014-0410-3>
- Kohler, H. P., Rodgers, J. L., & Christensen, K. (1999). Is fertility behavior in our genes? Findings from a Danish twin study. *Population and Development Review*, 25(2), 253–288.
- Knafo, A., & Plomin, R. (2006). Prosocial behavior from early to middle childhood: Genetic and environmental influences on stability and change. *Developmental Psychology*, 42(5), 771–786. <https://doi.org/10.1037/0012-1649.42.5.771>
- Kruger, D. J. (2017). Brief self-report scales assessing life history dimensions of mating and parenting effort. *Evolutionary Psychology*, 15(1). <https://doi.org/10.1177/1474704916673840>
- Lai, M. H., Richardson, G. B., & Mak, H. W. (2019). Quantifying the impact of partial measurement invariance in diagnostic research: An application to addiction research. *Addictive behaviors*, 94, 50–56.
- Li, W., Mai, X., & Liu, C. (2014). The default mode network and social understanding of others: what do brain connectivity studies tell us. *Frontiers in Human Neuroscience*, 8(74). <https://doi.org/10.3389/fnhum.2014.00074>
- Lowe, J. R., Edmundson, M., & Widiger, T. A. (2009). Assessment of dependency, agreeableness, and their relationship. *Psychological Assessment*, 21(4), 543–553.

- Mansolf, M., & Reise, S. P. (2017). When and why the second-order and bifactor models are distinguishable. *Intelligence*, *61*, 120–129.
- Maes, H. H., Silberg, J. L., Neale, M. C., & Eaves, L. J. (2007). Genetic and cultural transmission of antisocial behavior: an extended twin parent model. *Twin Research and Human Genetics*, *10*(1), 136–150. <https://doi.org/10.1375/twin.10.1.136>
- Mededović, J. (2018). Exploring the links between psychopathy and life history in a sample of college females: A behavioral ecological approach. *Evolutionary Psychological Science*, *4*(4), 466–473. <https://doi.org/10.1007/s40806-018-0157-5>
- Mededović, J. (2019). Harsh environment facilitates psychopathy's involvement in mating-parenting trade-off. *Personality and Individual Differences*, *139*, 235–240. <https://doi.org/10.1016/j.paid.2018.11.034>
- Muncer, S. (2011). The general factor of personality: Evaluating the evidence from meta-analysis, confirmatory factor analysis and evolutionary theory. *Personality and Individual Differences*, *51*, 775–778. <https://doi.org/10.1016/j.paid.2011.06.029>
- Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality*, *61*, 622–631.
- Mustanski, B., Viken, R. J., Kaprio, J., Winter, T., & Rose, R. J. (2007). Sexual behavior in young adulthood: A population-based twin study. *Health Psychology*, *26*(5), 610–617.
- Muthén, B., du Toit, S., & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Accessible at https://www.statmodel.com/download/Article_075.pdf
- Neale, M., & Cardon, L. (2013). *Methodology for genetic studies of twins and families* (Nato Science Series D, 67). Dordrecht: Springer Science & Business Media. (originally published in 1992).
- Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, *61*(6), 622–631.
- Olderbak, S., Gladden, P., Wolf, P. S. A., & Figueredo, A. J. (2014). Comparison of life history strategy measures. *Personality and Individual Differences*, *58*, 82–88.
- Penke, L., Denissen, J. A., & Miller, G. F. (2007). The evolutionary genetics of personality. *European Journal of Personality*, *21*, 549–587. <https://doi.org/10.1002/per.629>
- Polderman, T. J., Benyamin, B., De Leeuw, C. A., Sullivan, P. F., Van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*(7), 702–709.
- Richardson, G. B., Chen, C. C., Dai, C. L., Brubaker, M. D., & Nedelec, J. L. (2017c). The psychometrics of the Mini-K: Evidence from two college samples. *Evolutionary Psychology*, *15*(1). <https://doi.org/10.1177/1474704916682034>
- Richardson, G. B., Chen, C.-C., Dai, C.-L., Hardesty, P. H., & Swoboda, C. M. (2014). Life history strategy and young adult substance use. *Evolutionary Psychology*, *12*(5), 932–957.
- Richardson, G. B., Dai, C.-L., Chen, C.-C., Nedelec, J. L., Swoboda, C. M., & Chen, W.-W. (2016). Adolescent life history strategy in the intergenerational transmission and developmental stability of substance use. *Journal of Drug Issues*, *46*(2), 102–121
- Richardson, G. B., Dariotis, J. K., & Lai, M. H. (2017a). From environment to mating competition and Super-K in a predominantly urban sample of young adults. *Evolutionary Psychology*, *15*(1). <https://doi.org/10.1177/1474704916670165>
- Richardson, G. B., Sanning, B. K., Lai, M. H., Copping, L. T., Hardesty, P. H., & Kruger, D. J. (2017b). On the psychometric study of human life history strategies: State of the science and evidence of two independent dimensions. *Evolutionary Psychology*, *15*(1). <https://doi.org/10.1177/1474704916666840>
- Rodgers, J. L., Kohler, H. P., Kyvik, K. O., & Christensen, K. (2001). Behavior genetic modeling of human fertility: Findings from a contemporary Danish twin study. *Demography*, *38*(1), 29–42.
- Sear, R., & Mace, R. (2008). Who keeps children alive? A review of the effects of kin on child survival. *Evolution and Human Behavior*, *29*, 1–18.
- Snyder, H. R., Young, J. F., & Hankin, B. L. (2017). Strong homotypic continuity in common psychopathology-, internalizing-, and externalizing-specific factors over time in adolescents. *Clinical Psychological Science*, *5*(1), 98–110.
- Stearns, S. C. (1992). *The evolution of life histories*. Oxford University Press.
- Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance good enough? *Methodology*, *9*, 1–12.
- Stiver, K. A., & Alonzo, S. H. (2009). Parental and mating effort: Is there necessarily a trade-off? (Invited Review). *Ethology*, *115*(12), 1101–1126. <https://doi.org/10.1111/j.1439-0310.2009.01707.x>

- Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of man 1871–1971* (pp. 136–179). Aldine.
- van der Linden, D., Schermer, J. A., de Zeeuw, E., Dunkel, C. S., Pekaar, K. A., Bakker, A. B., & Petrides, K. V. (2018). Overlap between the general factor of personality and trait emotional intelligence: a genetic correlation study. *Behavior Genetics*, *48*(2), 147–154. <https://doi.org/10.1007/s10519-017-9885-8>
- van der Linden, D., Pekaar, K. A., Bakker, A. B., Schermer, J. A., Vernon, P. A., Dunkel, C. S., & Petrides, K. V. (2017). Overlap between the general factor of personality and emotional intelligence: A meta-analysis. *Psychological Bulletin*, *143*(1), 1–65.
- van der Linden, D., van Klaveren, D., & Dunkel, C. S. (2015). Emotional intelligence (EI) is an indicator of a slow life history strategy: A test of ability and trait EI. *Personality and Individual Differences*, *73*, 84–87.
- Verweij, K. J. H., Zietsch, B. P., Bailey, J. M., & Martin, N. G. (2009). Shared aetiology of risky sexual behaviour and adolescent misconduct: Genetic and environmental influences. *Genes, Brain and Behavior*, *8*(1), 107–113.
- Waldron, M., Heath, A. C., Turkheimer, E., Emery, R., Bucholz, K. K., Madden, P. A., & Martin, N. G. (2007). Age at first sexual intercourse and teenage pregnancy in Australian female twins. *Twin Research and Human Genetics*, *10*(3), 440–449.
- Wang, S., Chen, C. C., Dai, C. L., & Richardson, G. B. (2018). A call for, and beginner's guide to, measurement invariance testing in evolutionary psychology. *Evolutionary Psychological Science*, *4*, 166–178
- Wei, Y., De Lange, S. C., Scholtens, L. H., Watanabe, K., Ardesch, D. J., Jansen, P. R., ... & Posthuma, D. (2019). Genetic mapping and evolutionary analysis of human-expanded cognitive networks. *Nature Communications*, *10* (4839). <https://doi.org/10.1038/s41467-019-12764-8>
- West-Eberhard, M. J. (1979). Sexual selection, social competition, and evolution. *Proceedings of the American Philosophical Society*, *123*, 222–234
- Xu, T., Nenning, K. H., Schwartz, E., Hong, S. J., Vogelstein, J. T., Fair, D. A., ... & Langs, G. (2019). Cross-species functional alignment reveals evolutionary hierarchy within the connectome. *bioRxiv*, <https://doi.org/10.1101/692616>
- Zietsch, B. P., & Sidari, M. J. (2020). A critique of life history approaches to human trait covariation. *Evolution and Human Behavior*, *41*, 527–535. <https://doi.org/10.1016/j.evolhumbehav.2019.05.007>
- Zietsch, B. P., Verweij, K. J. H., Bailey, J. M., Wright, M. J., & Martin, N. G. (2010). Genetic and environmental influences on risky sexual behaviour and its relationship with personality. *Behavior Genetics*, *40*(1), 12–21.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.