The K-SF-42: A New Short Form of the Arizona Life History Battery

Aurelio José Figueredo[1], Rafael Antonio Garcia[1], J. Michael Menke[2],
W. Jake Jacobs[1], Paul Robert Gladden[3], JeanMarie Bianchi[4], Emily Anne Patch[1],
Connie J. A. Beck[1], Phillip S. Kavanagh[5], Marcela Sotomayor-Peterson[6],
Yunfan Jiang[7,8], and Norman P. Li[8]

Abstract

The purpose of the present article is to propose an alternative short form for the 199-item Arizona Life History Battery (ALHB), which we are calling the K-SF-42, as it contains 42 items as compared with the 20 items of the Mini-K, the short form that has been in greatest use for the past decade. These 42 items were selected from the ALHB, unlike those of the Mini-K, making direct comparisons of the relative psychometric performance of the two alternative short forms a valid and instructive exercise. A series of secondary data analyses were performed upon a recently completed five-nation cross-cultural survey, which was originally designed to assess the role of life history strategy in the etiology of interpersonal aggression. Only data from the ALHB that were collected in all five cross-cultural replications were used for the present analyses. The single immediate objective of this secondary data analysis was producing the K-SF-42 such that it would perform optimally across all five cultures sampled, and perhaps even generalize well to other modern industrial societies not currently sampled as a result of the geographic breadth of those included in the present study. A novel method, based on the use of the Cross-Sample Geometric Mean as a criterion for item selection, was used for generating such a cross-culturally valid short form.

The Psychometric Assessment of Life History (LH)

LH theory is a mid-level evolutionary theory governing optimal resource allocation among different components of fitness. The theory implicitly relies on regarding both the parent and the offspring as partially interchangeable vehicles for the ultimate function of genetic propagation. Selection upon the genome therefore allocates bioenergetic and material resources among those dedicated to parental survival (somatic effort) and those dedicated to offspring production and preservation (reproductive effort). Sexually reproducing organisms further partition reproductive effort among those resources needed for obtaining and retaining sexual partners (mating effort) and those needed for the longer term survivorship of the offspring produced (parental/nepotistic effort).

Although LH selection is ultimately based on genetic propagation, selection acts by shaping the proximate adaptive

[1] Department of Psychology, College of Science, University of Arizona, Tucson, AZ, USA
[2] A. T. Still Research Institute, A. T. Still University, Mesa, AZ, USA
[3] Department of Psychology, Sociology, and Criminal Justice, Middle Georgia State University, Macon, GA, USA
[4] Division of Integrated Sciences, Wilson College, Chambersburg, PA, USA
[5] School of Psychology, Social Work and Social Policy, University of South Australia, Adelaide, South Australia, Australia
[6] Department of Psychology and Communication, University of Sonora, Sonora, Mexico
[7] Singapore Prison Service, Singapore
[8] School of Social Sciences, Singapore Management University, Singapore

Corresponding Author:
Aurelio José Figueredo, Department of Psychology, School of Mind, Brain, and Behavior, College of Science, University of Arizona, 1503 East University Boulevard, PO Box 210068, Tucson, AZ 85721, USA.
Email: ajf@u.arizona.edu

profiles of resource allocations. Bioenergetic and material resources need to be mobilized to achieve the functional objectives of survival and reproduction. These resource mobilizations power the expression of the biological and behavioral adaptations that are the individual tactical elements deployed by the governing LH strategies.

The rationale that undergirds the psychometric approach to the assessment of LH strategies is to tap into the dynamics of these proximal resource allocations, rather than monitoring the attainment of the distal achievements to which they are directed. This is based upon the evolutionary psychological conception of organisms as adaptation executers rather than as fitness maximizers, where the resource allocations are the adaptations being executed and the distal achievements are the specific components of fitness being maximized. In contrast, the biodemographic approach directly samples the distal achievements eventually produced by these resource allocations, such as age of puberty, age of first sexual contact, number of sexual partners, age of first birth, interbirth interval, number of offspring at completed fertility, and so on.

Although the biodemographic approach has also proven useful in the study of LH strategies, it suffers from various limitations, as all single methods ultimately do. Among them is the well-known fallibility of the relation between *motivation* and *performance* in psychology. According to one formulation of this problem (Mitchell, 1982), performance can be modeled as the multiplicative product (or statistical interaction) of the individual's motivation, the individual's *ability*, and the individual's *environment*. This model states explicitly that, aside from the individual's motivation, the individual requires the physical and mental ability to perform the task as well as an environment that affords the requisite resources for the goal's attainment.

Applying this specifically to LH strategies, a human adolescent male's age of first sexual contact and number of sexual partners are not uniquely determined by his motivational state, which in most cases we can safely presume to be quite high. It is also codetermined by a number of intrinsic individual characteristics, such as his sexual attractiveness, social skills, and social status, and additionally by a number of extrinsic ecological characteristics, such as the availability of willing female sexual partners, the degree of direct access to such willing female sexual partners afforded by the environment, and the socially imposed restrictions (or lack thereof) governing adolescent sexual activity. A human adolescent male's age of first sexual contact and number of sexual partners are thus not very pure measures of his either slower of faster LH strategy, which is a highly heritable trait and thus not very contingent upon such transient ambient circumstances ($h^2 \sim .65$; Figueredo & Rushton, 2009; Figueredo, Vásquez, Brumbach, & Schneider, 2004). Similarly, a human female's age of first birth, interbirth intervals, and number of offspring at completed fertility is codetermined by a number of individual difference factors, such as her physical condition and a number of ecological and sociocultural factors, such as environmental resource abundance and any socially imposed regulation of either the timing

or the relational context of reproductive activity. At the present time, we believe that there is insufficient empirical evidence reported in the published literature to determine whether there might or might not be any sex differences in the heritability of LH strategy.

This is not to say that the psychometric approach is somehow more valid than the biodemographic approach or vice versa. The idea is that they are *complementary*, as they each provide windows into different successive stages of the same causal processes (Garcia et al., 2016). The psychometric approach is based on sampling *behavioral and cognitive indicators* of the LH resource allocations among different component domains of fitness. For example, the Arizona Life History Battery (ALHB) contains seven subscales (not including the Mini-K, which is designed as domain general) that sample the following domain-specific psychosocial resource allocations: (1) *Insight, Planning, and Control*; (2) *Mother/Father Relationship Quality*; (3) *Family Social Contact and Support*; (4) *Friends Social Contact and Support*; (5) *Romantic Partner Attachment*; (6) *General Altruism*; and (7) *Religiosity*. This sampling of resource allocation domains is not represented as completely exhaustive. However, we believe that it provides reasonably adequate coverage of the major areas of psychosocial resource investment that LH theory would predict as characterizing a *slower* LH strategy; *faster* LH strategists are thus predicted to score *lower* on the latent common factor that can be extracted from these seven more domain-specific subscales.

The Mini-K and the K-SF-42

Since it was first published 10 years ago (Figueredo et al., 2006), the Mini-K has been a highly successful 20-item short form for the psychometric measurement of LH strategy. It has fared very well in meta-analyses (e.g., Figueredo et al., 2014), with a mean bivariate correlation of $\bar{r} = .75$ and a disattenuated synthetic correlation of $\rho = .91$ with its corresponding long form, the 199-item ALHB. Such meta-analytic work, however, has been limited to English-speaking North American undergraduate student populations. The concern is that the Mini-K has actually been applied to a much wider variety of populations cross-culturally including such 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 & Hoben, 2013; Buunk, Pollet, Klavina, Figueredo, & Dijkstra, 2009; Cabeza de Baca, Figueredo, & Ellis, 2012; Cabeza de Baca, Sotomayor-Peterson, Smith-Castro, & Figueredo, 2014; Egan et al., 2005; Figueredo, Andrzejczak, Jones, Smith-Castro, & Montero-Rojas, 2011; Figueredo, Tal, McNeill, & Guillén, 2004; Figueredo & Wolf, 2009; Frías-Armenta et al., 2005; Frías-Armenta, Valdez-Ramirez, Nava-Cruz, Figueredo, & Corral-Verdugo, 2010; Gaxiola-Romero & Frías-Armenta, 2008; Jonason, Li, & Czarna, 2011; Sotomayor-Peterson, Cabeza de Baca, Figueredo, & Smith-Castro, 2013; Sotomayor-Peterson, Figueredo, Hendrickson-Christensen, & Taylor, 2012; Tal, Hill, Figueredo, Frías-Armenta, & Corral-Verdugo, 2006; Tifferet, Agrest, &

Shlomo Benisti, 2011; van der Linden, Figueredo, de Leeuw, Scholte, & Engels, 2012; Woodley of Menie & Madison, 2015).

An acknowledged disadvantage of the Mini-K is its limited reliability ($\bar{\alpha} = .75$ in the cited meta-analysis), which is due the fact that it is a short form for an entire multivariate latent variable (the K-Factor), comprising several component sub-scales, rather than a single uniform scale. The latter fact inevitably leads to some irreducible degree of heterogeneity among the items, as they were designed to reflect conceptually distinct but convergent indicators of the multivariate latent construct. A measure of internal consistency among items, such as Cronbach's α, will therefore return an appropriately low estimate. Also worth noting is that the test–retest reliability of the Mini-K is similarly low, which can only be attributed to the relatively low number of items.

The 20 items on the Mini-K are not a subset of the 199 items on the ALHB but are instead unique items written to assess its content areas more globally. For example, 2 items ("While growing up, I had a close and warm relationship with my biological mother" and "While growing up, I had a close and warm relationship with my biological father") were written to summarize the substantive content of the 26-item Mother/Father Relationship Quality subscale (adapted from Brim et al., 2000; Figueredo, Vásquez, Brumbach, & Schneider, 2004, 2007) of the ALHB. This strategy was deemed optimal for producing a 20-item short form in comparison with selecting the "best" items from the subscale, as even the 2 best-performing individual items (by whatever criterion) were unlikely to cover the entire substantive content of the subscale and reflect the full breadth of the construct. The Mini-K was originally developed for use in situations wherein time constraints made use of the full ALHB prohibitive. It was never recommended for situations where the psychometric assessment of LH strategy was intended as a cornerstone of the research; in such cases, the full ALHB was to be preferred and was presumed to be worth the cost in added response burden. Additionally, the recent completion of a large cross-cultural survey using the ALHB (Figueredo et al., 2016) presents an opportunity to create a short form that can also stand up to tests of cross-cultural validity. We illustrate a novel method for generating such a cross-culturally valid short form.

Description of the Present Study

These characteristics of the Mini-K offer an opportunity, after a decade of productive use, to propose a somewhat longer short form that might retain some of its advantages, avoid its major disadvantages, and allow for more flexibility in research application. The purpose of the present article is to propose such an alternative short form for the ALHB, which we are calling the K-SF-42, as it contains 42 items (see Appendix for a complete listing). These 42 items are directly selected from the ALHB, unlike the 20 items of the Mini-K, making direct comparisons of the relative psychometric performance of the two alternative short forms a potentially valid and instructive exercise.

The purpose of this article is not, however, to present a review of the construct validity of the ALHB itself, as we have already published a rather extensive meta-analysis on both the *predictive* ("nomological") and *convergent* validity of the Mini-K and the ALHB (Figueredo et al., 2014). We have also published a more complete treatment of the *construct* validity of the hierarchical latent structure of the Mini-K, the ALHB, and related measures using multiple data sets and applying confirmatory factor analyses as well as other methods (Olderbak, Gladden, Wolf, & Figueredo, 2014). Furthermore, we have also expounded eloquently elsewhere upon the epistemological and methodological rationales underlying the psychometric approach to assessing LH strategy (Figueredo, Cabeza de Baca, & Woodley, 2013; Figueredo et al., 2015).

To maintain the breadth of the construct at each level of aggregation, 6 items were selected from each of the seven subscales of the ALHB (not including the Mini-K), representing several domains of potential allocation of bioenergetic and material resources in a manner consistent with the execution of a slow LH strategy. Where appropriate, these 6 items were balanced across the multiple subdomains sampled by each subscale. As novel item selection procedures were applied to obtain items and subscales that functioned at the highest levels of cross-cultural generality, we followed up with item response theory (IRT) Rasch modeling (Bond & Fox, 2001) of both individual items and aggregated subscales to ascertain whether our procedures performed as intended. We also performed a generalizability theory (GT) analysis (Figueredo & Olderbak, 2008) on the IRT-generated subscale difficulty estimates to test for the degree of cross-cultural invariance that our procedures managed to achieve.

This series of secondary data analyses were performed upon a recently completed five-nation cross-cultural survey, which was originally designed to assess the role of LH strategy in the etiology of interpersonal aggression (Figueredo et al., 2016). Only data from the ALHB, collected in all five cross-cultural replications, were used for the present analyses. The single immediate objective was producing the K-SF-42, such that it would perform optimally across all five cultures sampled, and perhaps even generalize well to other cultures not currently sampled as a result of the geographic breadth of those included in the present study.

Method

Participants and Procedures

A total of 738 undergraduate students participated in the study, enrolled in introductory psychology courses in five different major universities located in Australia (AU; $N = 131$), Italy (IT; $N = 172$), Mexico (MX; $N = 160$), Singapore (SG; $N = 115$), and the United States (US; $N = 160$). In the original study from which these data are derived (Figueredo et al., 2016), participants completed a series of computerized self-report questionnaires assessing the following constructs: LH strategies; executive functioning; mating strategies; mate values; mating

effort; revenge ideologies; psychopathic and aggressive attitudes; and their coercive, aggressive, and possibly violent behavioral interactions over the past year with any and all members of their same sex and with any and all members of the opposite sex (whether or not they were intimate romantic partners) with which they might have interacted. Participants signed up for the study, provided informed consent, and completed the questionnaires through a secured Internet website.

Measures

The ALHB. The ALHB (Figueredo et al., 2007) is a set of cognitive and behavioral indicators of LH strategy compiled and adapted from various original sources. These self-report psychometric indicators measure graded individual differences along various complementary facets of a coherent and coordinated LH strategy, as specified by LH theory, and converge upon a single multivariate latent construct. These were scored and aggregated directionally to indicate a slow (high-*K*) LH strategy, on the fast–slow (*r*–*K*) continuum. Not including the Mini-K, the seven subscales of the ALHB are as follows: Insight, Planning, and Control (20 items; adapted from Brim et al., 2000; Figueredo et al., 2004, 2007); Mother/Father Relationship Quality (26 items; adapted from Brim et al., 2000; Figueredo et al., 2004, 2007); Family Social Contact and Support (15 items; adapted from Barrera, Sandler, & Ramsay, 1981; Figueredo et al., 2001); Friends Social Contact and Support (15 items; adapted from Barrera et al., 1981; Figueredo et al., 2001); Experiences in Close Relationships (ECR: 36 items; adapted from Brennan, Clark, & Shaver, 1998; Alonso-Arbiol, Balluerka, & Shaver, 2007; Figueredo et al., 2005); General Altruism (50 items; adapted from Brim et al., 2000; Figueredo et al., 2004, 2007); and Religiosity (17 items; adapted from Brim et al., 2000; Figueredo et al., 2004, 2007). As the internal consistency reliabilities of the ALHB subscales are a major subject of the present analyses, presentation of these psychometric properties will be deferred to the Results section.

The Mini-K. The Mini-K (Figueredo et al., 2006) is a 20-item short form of the ALHB. Meta-analytically (e.g., Figueredo et al., 2014), the Mini-K has shown a high mean bivariate correlation ($\bar{r} = .75$) and a higher disattenuated synthetic correlation of ($\rho = .91$) with its corresponding 199-item long form. The items on the Mini-K are not directly sampled from the seven other subscales comprising the ALHB, but are instead designed to provide more global assessments within each of the LH domain-specific resource allocations sampled by the ALHB. As with the ALHB, the Mini-K is scored directionally to indicate a slow (highly K-selected) LH strategy on the "fast–slow" (*r*–*K*) continuum.

Statistical Analyses

All statistical analyses were performed using MS Excel 2003 (Microsoft, 2003), SAS 9.3 software (SAS Institute Inc., 2011), and WINSTEPS 3.68.1 (Linacre, 2016).

Certain theoretically motivated decisions were made at the beginning of this process regarding items that would not be the candidates for inclusion in the new short form, the K-SF-42. One such decision was to eliminate from consideration the set of items in the General Altruism scale sampling Altruism Toward Children, as it tended to produce an overabundance of missing data in younger samples of participants, and was not deemed theoretically essential to the construct being measured.

The cross-sample geometric mean (CSGM) method for optimal item selection. To maintain the breadth of the construct at each level of aggregation, 6 items were selected from each of the seven subscales of the ALHB (not including the Mini-K), representing several domains of potential allocation of bioenergetic and material resources in a manner consistent with the execution of a slow LH strategy. Where appropriate, these 6 items were balanced across the multiple subdomains sampled by each subscale. For the Mother/Father Relationship Quality subscale, the 6 items were selected by sampling the best 3 items from each of the parental (Mother and Father) subdomains. Similarly, for the General Altruism scale, the 6 items were selected by sampling the best 2 items from each of the following three subdomains: Altruism Toward Kin, Altruism Toward Friends, and Altruism Toward Community.

The way that we operationalized which items were "the best" defined the novelty of this approach and has the beauty of its great simplicity. We first calculated the part-whole correlation of each item score with the unit-weighted (UW) factor score for the subscale in which it was nested, conceptually representing the UW factor loadings of each item. We did this separately for each of the five cross-cultural replications and pasted these as columns on a single MS Excel spreadsheet. We used the spreadsheet to compute the geometric mean of these part-whole correlations across all five cross-cultural replications. Next, we selected the 6 items with the highest geometric means from each subscale (within the subdomain constraints outlined above) for inclusion in the K-SF-42, by simply using the spreadsheet to sort the data by the columns containing the subscales and the geometric means in descending order.

The rationale for using the geometric rather than arithmetic mean is that it involves the multiplication rather than the addition of the parameter estimates for each of the five cross-cultural replications. In the case of the arithmetic mean, the sum is divided by *n*; in the case of the geometric mean, the *n*th root of the product is instead calculated. Multiplication is equivalent in Boolean algebra to a logical "AND" statement or logical conjunction; whereas addition is equivalent in Boolean algebra to a logical "OR" statement or logical disjunction. A high product term is thus only produced when the values in *all* the coefficients being multiplied are high; a high additive term, on the other hand, can still be obtained if one or more of the coefficients being summed are sufficiently high to *compensate* for any that might be lower. Thus, a higher geometric mean is a stronger filter than a high arithmetic mean for selecting items that perform sufficiently well across *all* (rather than mere *some*)

of the multiple cultures sampled. We have therefore named this the CSGM method for optimal item selection, having been first pioneered in a previous study but gone without any special designation (Patch, Garcia, Figueredo, & Kavanagh, 2016).

In contrast, other short form construction methods that rely on empirical selection of variables will often experience problems of replication when applied to independent samples, especially cross-cultural ones. Part of this problem arises from model fitting to sample-specific characteristics that are not present in other samples. For example, empirical selection methods may capitalize on chance differences (stochastic fluctuations) among item characteristics that are purely due to sampling error; as a result which items are truly the best may not replicate between one independent and another even if they are drawn from the same population, being analogous to statistical "Type I Errors." Furthermore, empirically based selection of locally optimal items can lead to disastrous results when applied outside of the population upon which the new scale was originally developed. Although this latter circumstance is obviously less of a problem the more alike the samples are to each other, between-sample differences on relevant characteristics can be quite large in the case of cross-cultural research.

After implementing the CSGM procedures, we had to assess whether this seemingly clever tactic actually worked in practice by performing each of the following psychometric assessment procedures: (1) We compared the internal consistency reliabilities of the short form with the long form for each of the seven subscales of the ALHB, (2) we compared the part-whole correlations (UW factor structures) of the short form with the long form for each of the seven subscales of the ALHB, (3) we examined the simple bivariate correlations of the short form with the long form for each of the seven subscales of the ALHB, (4) we examined the relative internal consistency reliabilities of the two alternative short forms (Mini-K and K-SF-42) of the ALHB, each as a unitary scale irrespective of internal structure, as well as the correlations of those UW scale scores with the UW factor score of the full ALHB, and (5) we examined the relative incremental validities of the two alternative short forms (Mini-K and K-SF-42) of the ALHB in multiple regression models predicting the UW factor score of the full ALHB.

IRT Rasch modeling of item and subscale difficulties. To validate the results obtained by means of this novel method by more conventional means, we ran a one parameter Rasch model to obtain IRT parameter estimates on the K-SF-42 items and subscales. The IRT parameter of primary interest was the so-called item difficulty, which indicates how high a factor score on the underlying latent variable (θ) is required to produce a "correct" response on any given item (Bond & Fox, 2015). The latent variable of interest is typically cognitive ability (g), but in the present case it is simply LH speed (K). The IRT method chosen was the single parameter Rasch model and analyzed with WINSTEPS. First, countries were combined within subscales, and each subscale was analyzed with countries as items. This was done as a way to see whether countries differed in their relative

rank within each subscale, a rough indication of differential item functioning of countries within subscales. Second, subscale totals were computed and then designated as items in an overall total scale score, also in a Rasch analysis. This allowed for subscales to be ranked in terms of "difficulty" in terms of their ease of endorsement in their overall LH strategy. The easiest subscales would be those with the highest priority of endorsement—a strategy most commonly shared in the population. After computing subscale difficulties by country, they could be inspected for relative difficulty or relative importance to each population.

The rationale behind these analyses was that it was important to determine whether our novel item selection procedure, based on classical test theory (CTT), might have selected items that were too similar on item difficulty as a result of being highly correlated with each other. CTT item selection procedures have been criticized for having this bias (Bond & Fox, 2015), as a scale composed of items of similar difficulties will not adequately measure individual persons throughout the extent of the latent trait, and thus may incorrectly assess the latent trait score of persons either substantially above or substantially below that central tendency. In the case of the K-SF-42, we had endeavored to preserve the autonomy among these subscales by the hierarchical design of our item selection procedure, as we understood that limitation in advance and took that step to hopefully avoid it. That is why the CTT procedure that we used only selected the best (most highly correlated) items *within* each subscale. We hypothesized that although the items within subscales might have similar item difficulties, each of the aggregated 6-item subscales should nonetheless have different difficulties. Assessing difficulties was in this case assessing their range of sensitivity of each of the subscales to the latent construct of LH speed (K). This latter prediction was based on the premise underlying the development of the ALHB, where each of the subscales represents a distinct domain-specific opportunity for an LH-based resource allocation decision to differentially invest (or not) in that particular system of social relationships. As each specific resource allocation domain should have a different item difficulty in terms of LH speed, we should therefore be able to observe substantial variance among the subscale difficulties. The hierarchical item selection procedure that we used should therefore enable the K-SF-42 to measure the full range of LH speed (K) along the broader difficulty distribution of subscales.

Generalizability analyses of item and subscale difficulties. The other important question that we hoped to address by these analyses was whether the subscale difficulties differed substantially across the sampled cultures. As we did not have enough data in each sample for a formal differential item functioning analysis (Acar, 2011; Linacre, 2013, 2016; Scott et al., 2009; Tristan, 2006), we instead subjected the individual-level subscale difficulty estimates to general linear modeling (GLM) and GT analyses (Figueredo & Olderbak, 2008). We restricted these analyses to aggregate subscale scores, as the cross-cultural

Table 1. Internal Consistency Reliabilities for Convergent Indicators of Life History Strategy in Australia (AU), Italy (IT), Mexico (MX), Singapore (SG), and the United States (US).

Scale/Subscale	Long Form Cronbach's					Short Form Cronbach's				
	AU	IT	MX	SG	US	AU	IT	MX	SG	US
The Arizona Life History Battery										
Insight, Planning, and Control	.91	.85	.95	.88	.89	.86	.79	.92	.82	.86
Parental Relationship Quality	.90	.90	.94	.85	.90	.83	.83	.85	.80	.81
Family Contact and Support	.91	.92	.89	.94	.92	.92	.90	.91	.93	.92
Friends Contact and Support	.90	.88	.90	.86	.90	.89	.87	.89	.90	.90
Romantic Partner Attachment	.93	.92	.89	.92	.91	.83	.81	.84	.85	.81
General Altruism	.92	.90	.94	.93	.90	.77	.75	.77	.78	.71
Religiosity	.95	.95	.96	.96	.95	.91	.93	.94	.96	.94

Table 2. Unit-Weighted (UW) Factor Structures for Convergent Indicators of Life History Strategy in Australia (AU), Italy (IT), Mexico (MX), Singapore (SG), and the United States (US).

Scale/Subscale	Long Form Part-Whole Correlations (UW λ)*					Short Form Part-Whole Correlations (UW λ)*				
	AU	IT	MX	SG	US	AU	IT	MX	SG	US
The Arizona Life History Battery										
Insight, Planning, and Control	.64	.58	.77	.54	.63	.56	.58	.69	.56	.60
Parental Relationship Quality	.60	.61	.61	.70	.49	.55	.55	.66	.62	.45
Family Contact and Support	.60	.60	.59	.52	.62	.62	.61	.65	.49	.65
Friends Contact and Support	.40	.50	.61	.43	.58	.45	.51	.53	.45	.58
Romantic Partner Attachment	.40	.44	.39	.30	.38	.38	.43	.36	.33	.41
General Altruism	.70	.72	.68	.57	.66	.63	.62	.61	.54	.62
Religiosity	.22	.44	.58	.37	.47	.29	.46	.60	.47	.49

*All part-whole correlations (UW factor loadings) are statistically significant at $p < .05$.

functioning of the individual items within subscales was of lesser interest, for reasons already stated.

A GLM was constructed, analogous to a two-way analysis of variance, where we tested for effects on the person-level difficulty estimates of: (1) the different cross-cultural samples, (2) the different subscales, and (3) the two-way interaction between them. We then followed up with estimated mean squares and variance components analysis using the same design to obtain estimates of the latent variance components hypothetically underlying these observed variances, first using Type 1 and then using reweighted maximum likelihood (REML) estimation for comparison. Finally, we used these variance component estimates to calculate generalizability coefficients for each of the two random effects of interest with respect to their interaction with each other, based on the Type 1 and REML estimates obtained.

Results

Table 1 presents side-by-side comparisons of the internal consistencies of the K-SF-42 and the full ALHB for all five cross-cultural sites. All Cronbach's α for both the K-SF-42 and the full ALHB met the criterion for "acceptable" (α ≥ .70) with 64 of 70 meeting the criterion for "good" (α ≥ .80) and 40 of 70 meeting the criterion for "excellent" (α ≥ .90). Across all five cross-cultural sites, only one scale (General Altruism) demonstrated an appreciable decrease in internal reliability when

going from the full ALHB to the K-SF-42 (mean Δα = −.16). This difference, however, did not drop the value of Cronbach's α below the acceptable threshold (range α = .71–.78, median = .77). The other six scales showed a decrease of no more than .11 in Cronbach's α when using the K-SF-42 (range Δα = −.11 to .04, median Δα = −.03).

Table 2 presents the UW factor structures of the K-SF-42 and the full ALHB for all five cross-cultural sites. Only two part-whole correlations showed a change of greater than .10 when going from the full ALHB to the K-SF-42, Religiosity in SG (ΔUW λ = +.10) and General Altruism in IT (ΔUW λ = −.10). Nevertheless, these changes do not appear to decrease the part-whole correlations into a range of questionable validity. The statistical nonsignificance of average scale part-whole change between forms seems to support this (AU mean ΔUW λ = −.01, p > .05; IT mean ΔUW λ = −.02, p > .05; MX mean ΔUW λ = −.02, p > .05; SG mean ΔUW λ = .00, p > .05; US mean ΔUW λ = −.00), p > .05).

Table 3 presents the short–long correlations between the K-SF-42 and the full ALHB for all five cross-cultural sites. All short–long correlations are strong and positive (range = .74–.98, median = .90), indicating a high degree of agreement between the short and long form when measuring each domain-specific component of slow LH.

Table 4 presents side-by-side comparisons of the internal consistencies and short–long form correlations of the Mini-K Short Form and the K-SF-42 for all five cross-cultural sites. On

Table 3. Short-Long Form Correlations for Convergent Indicators of Life History Strategy in Australia (AU), Italy (IT), Mexico (MX), Singapore (SG), and the United States (US).

Scale/Subscale	Short-Long Form Correlations*				
	AU	IT	MX	SG	US
The Arizona Life History Battery					
Insight, Planning, and Control	.91	.81	.92	.83	.87
Parental Relationship Quality	.89	.86	.90	.78	.90
Family Contact and Support	.91	.93	.86	.95	.91
Friends Contact and Support	.93	.91	.87	.90	.89
Romantic Partner Attachment	.93	.86	.90	.92	.89
General Altruism	.85	.83	.87	.88	.88
Religiosity	.98	.97	.98	.98	.98

*All short-long form correlations are statistically significant at $p < .05$.

all metrics and for all sites, the K-SF-42 outperformed the Mini-K Short Form. Cronbach's α comparison tests showed a statistically significant improvement of the K-SF-42 over the Mini-K Short Form, AU $\chi^2(1) = 8.74$, $p < .05$; IT $\chi^2(1) = 30.30$, $p < .05$; MX $\chi^2(1) = 3.94$, $p < .05$; SG $\chi^2(1) = 13.64$, $p < .05$; US $\chi^2(1) = 25.79$, $p < .05$, by a median of .12 (range = [.04, .18]). The same was true for short–long correlation comparison tests, AU $t(130) = -4.81$, $p < .05$; IT $t(171) = -4.22$, $p < .05$; MX $t(159) = -8.4$, $p < .05$; SG $t(115) = -6.18$, $p < .05$; US $t(159) = -6.97$, $p < .05$, which showed improvement by a median of .13 (range = [.10, .18]).

Table 5 presents the incremental validities of the Mini-K Short Form and the K-SF-42 for all five cross-cultural sites. When both short forms were entered into a regression to predict the full ALHB, the total explained variance was high (R^2 range = range = [.89, .96] median = .93), indicating that the two short forms taken together capture most of the variance of the full ALHB. The partial R^2 for each short form represents the proportion of the variance that is uniquely explained by that short form. In other words, it is the incremental validity of that short form over the other. All partial R^2s were statistically significant; however, the partial R^2s of the K-SF-42 were approximately between 3.5 and 11.5 times greater than those of the Mini-K Short Form (median = 4 times). This means that while the Mini-K Short Form and the K-SF-42 captured much of the same variance of the full ALHB, the K-SF-42 alone accounted for more uniquely; it had greater incremental validity.

Column 1 of Table 6 displays the standard deviations of the raw item difficulties within subscales. We believe that these were adequate dispersions, with the possible exception of the General Altruism subscale. Although we do not know why the dispersion among items for that one subscale was lower than the others, we believe that these standard deviations generally indicate a reasonably adequate dispersion of the items within subscales. Columns 2–5 of Table 6 display the standardized subscale Infit/Outfit Mean Squares and standardized subscale Infit/Outfit Standard Deviations for each of the subscale scores. For efficiency of presentation, subscale scores were aggregated across all five cross-cultural samples. We believe these to be

excellent indices of both infit and outfit for all of the subscales, indicating an adequate data–model fit.

Table 7 displays the results of our generalizability analyses of the IRT-based subscale difficulties for the K-SF-42. We did not estimate the generalizability across cultures of individual item scores as these were of lesser interest given that using the CSGM method as a criterion for item selection does not guarantee preserving optimal ranges of difficulties among items within subscales.

First, the preliminary GLM shows that the preponderance of the systematic variance in difficulties is accounted for by the main effect of subscale (73.9%), supporting the greater variance among subscales we predicted, with only relatively trivial amounts accounted for by the main effect of Sample (7.9%) and the Sample × Subscale interaction (18.1%), indicating that the variance in difficulties among the different cross-cultural samples was not very substantial and that the variance in the rank ordering of subscale difficulties different among cross-cultural samples was also relatively minor.

Second, we performed an expected mean squares analysis, modeling both Sample and Subscale as random factors, such that the interaction term (Sample × Subscale) intruded on each of them and thus served as the common denominator term for both main effects in all further calculations. This permitted the estimation of the hypothetical variance components by two convergent methods, hierarchical Type 1 sums of squares (conservatively assigning causal priority to Sample) and REML, and these are tabulated for comparison within the last two columns.

Third, these estimated variance components were used to compute the following two generalizability coefficients, using each of these two estimation methods: (1) for Sample as the focal facet, being generalized across the Sample × Subscale interaction as the random facet and (2) for Subscale as the focal facet, being generalized across the Sample × Subscale interaction as the random facet. These indicated that the variance in difficulties among cross-cultural samples did not generalize very well across the subscales ($E^2_{Type1} = .186$; $E^2_{REML} = .186$), whereas indicated that the variance in difficulties among the subscales did generalize very well across the cross-cultural samples ($E^2_{Type1} = .753$; $E^2_{REML} = .753$). These results were identical to three decimal places across estimation methods and entirely in line with our predictions, supporting our claim that the relative difficulties of the subscales are quite consistent across cultures, indicating that the subscales function in similar ways in measuring the construct of interest.

As a visual illustration of these results, Figure 1 shows a consistency in the relative difficulty of the seven subscales across the five countries, a consistency observed in the "all" countries category. The one apparent inconsistency is the family scale difficulty for the Singapore sample. Specifically, the Singapore family score was rated as very easy in comparison to the other four countries. This especially low score on that subscale, as compared with the other cross-cultural samples, may be attributable to traditional Asian family values ultimately based on the "filial piety" preached by the Eastern Sage Confucius (孔夫子 Kǒng Fūzǐ, 551–479 BC; Soothill, [1910] 1937).

Table 4. Internal Consistency Reliabilities of Alternative Short Forms (Mini-K and K-SF-42) for Latent Multivariate Construct of Life History Strategy in Australia (AU), Italy (IT), Mexico (MX), Singapore (SG), and the United States of America (US).

	Alternative Short Form Cronbach's					Alternative Short-Long Form Correlations*				
Scale/Subscale	AU	IT	MX	SG	US	AU	IT	MX	SG	US
The Arizona Life History Battery										
Mini-K Short Form (20 Items)	.76	.66	.85	.73	.72	.83	.81	.78	.76	.81
K-SF-42 (Standardized Items)†	.84	.84	.89	.85	.86	.93	.91	.96	.93	.94

*All short-long form correlations are statistically significant at $p < .05$. †All Cronbach's and short–long correlations for the K-SF-42 are statistically greater than those of the Mini-K at $p < .05$.

Table 5. Incremental Validities of Alternative Short Forms (Mini-K and K-SF-42) for Latent Multivariate Construct of Life History Strategy in Australia (AU), Italy (IT), Mexico (MX), Singapore (SG), and the United States (US).

		pR^2	
Site	Total R^2	Mini-K Short Form	K-SF-42
The Arizona Life History Battery			
AU	.93	.06	.25
IT	.89	.07	.25
MX	.96	.03	.35
SG	.91	.05	.33
US	.93	.05	.28

Note. All partial correlations are statistically significant at $p < .05$.

Discussion

The new and slightly longer short form of the ALHB (as compared with the Mini-K), which we named the K-SF-42, performed very well with respect to various psychometric standards against which it was evaluated. The K-SF-42 reduced the number of items needed to assess slower LH from 199 items in the original ALHB to only 42 items in the new short form, and it did so with minimal losses of internal consistency reliability and convergent validity among the constituent subscales. We believe that these psychometric indices of performance would have been deemed adequate had we submitted the original interpersonal aggression study for publication using the K-SF-42 in place of the full ALHB; it is doubtful that any reasonable editor or peer reviewer would have been concerned with the adequacy of reported psychometric benchmarks had they not seen the slightly better ones provided by the full ALHB.

In addition, we found high correlations between the UW factor scores for the reduced K-SF-42 subscales and the full ALHB subscales, between the UW factor scores for the overall K-SF-42 and ALHB, and even between the UW factor scores for the overall K-SF-42 and Mini-K. Treating the K-SF-42 as a unitary scale, irrespective of internal subscale structure, we found small but statistically significant improvements in the K-SF-42 with respect to the Mini-K in their internal consistency reliabilities and predictive validity coefficients in relation to the full ALHB. Testing for incremental validity of the K-SF-42 with respect to the Mini-K in predicting the full ALHB

score, we found their overlap to be quite substantial, but the unique proportions of variance predicted by the K-SF-42 to be substantially greater in magnitude predicted than those predicted by the Mini-K, although these latter variance components were also statistically significant.

We therefore conclude that the K-SF-42 has been developed such that it performs optimally across all five cultures sampled, and that these results will perhaps prove to generalize similarly to most other modern industrial societies that were not currently sampled as a result of the geographic breadth of those included in the present study. We would therefore recommend the K-SF-42 for use over the Mini-K for situations in which the total acceptable response burden of the study can bear the weight of more than 20 items but less than 199 items for the psychometric assessment of LH strategy.

As the K-SF-42 has been newly introduced with the present article, there are as yet no data on its predictive validity. Nevertheless, we believe that the magnitudes of the reported correlations of the K-SF-42 with both the Mini-K and the full ALHB are sufficiently high that we can confidently expect that its predictive validities would be comparable to the latter, lying somewhere between that of the Mini-K and the full ALHB, and moderated only by its intermediate reliability. As cited above, the meta-analytically disattenuated correlation coefficient (ρ) between the Mini-K and the full ALHB was .91 (Figueredo et al., 2014); although we do not yet have sufficient data on the K-SF-42 to make such a precise estimate, there is not reason to believe that it would not be even higher, given the evidence that we presented in the present article. Under that assumption, it is a straightforward logical inference to predict that the K-SF-42 should have stronger predictive power than the Mini-K for all criterion variables that have been assessed so far.

Finally, we have validated a novel methodology for producing cross-culturally valid short forms of longer measures using the CSGM method for optimal item selection. We have also cross-validated this novel method using IRT Rasch analyses of the item and subscale difficulty levels, followed by GT analyses that showed acceptable levels of consistency in the rank ordering of aggregate subscale difficulties across the cultures sampled. Thus, we have shown that the CSGM method can produce psychometrically acceptable and robust results, at least in application to the present data.

We understand that we have not conformed to the conventional "best practices" psychometric standards in these

Table 6. Standard Deviations (SD) of Raw Item Difficulties Within Standardized Subscale Infit/Outfit Mean Squares (MNSQ) and Standardized Subscale Infit/Outfit SD for Convergent Indicators of Life History Strategy.

Subscales	SD of Item Difficulties Within Subscales	Subscale Z-Infit MNSQ	Subscale Z-Infit SD	Subscale Z-Outfit MNSQ	Subscale Z-Outfit SD
Insight, Planning, and Control	1.45	0.78	−3.68	0.76	−3.91
Parental Relationship Quality	1.48	0.84	−2.63	0.84	−2.56
Family Contact and Support	2.65	0.85	−3.34	0.85	−3.05
Friends Contact and Support	2.03	0.89	−1.95	0.88	−1.99
Romantic Partner Attachment	1.09	1.10	1.75	1.13	2.21
General Altruism	0.64	0.79	−3.97	0.78	−4.06
Religiosity	1.67	0.94	−1.04	0.97	−0.53

Table 7. Generalizability Analysis of IRT-Based Subscale Difficulties for the K-SF-42.

General Linear Model (Observed Variances)

Source	DF	Semipartial η^2	F-Ratio	p (Ho)
Sample	4	.079	2.60*	.061
Subscale	6	.739	16.26*	<.0001
Sample × Subscale	24	.182		

Expected Mean Squares Analysis and Estimated Variance Components Analysis

Source	DF	Expected Mean Squares	Type 1 Estimates	REML Estimates
Sample	4	σ^2 (Error) + σ^2 (Sample × Subscale) + 7 σ^2(Sample)	.071	.071
Subscale	6	σ^2 (Error) + σ^2 (Sample × Subscale) + 5 σ^2 (Subscale)	.941	.941
Sample × Subscale	24	σ^2 (Error) + σ^2 (Sample × Subscale)	.309	.309

Generalizability Coefficients

Focal Facet	Random Facet	Type 1 E^2 Estimates	REML E^2 Estimates
σ^2(Sample)	σ^2 (Sample × Subscale)	.186	.186
σ^2(Subscale)	σ^2 (Sample × Subscale)	.753	.753

Note. IRT = item response theory; REML = reweighted maximum likelihood.
*Statistically significant at $p < .05$.

empirical demonstrations, mostly due to limitations in the sample size available from each cross-cultural replication, but we emphasize that one of our major objectives in writing this article was to introduce a novel method for creating short forms out of long forms by using data from multiple independent samples in a creative new way. Nevertheless, we have compared the performance of this novel method with those of a more conventional one, and we interpret the results of the IRT Rasch analyses and associated GT analyses as generally supporting the efficacy of the CSGM method for creating the K-SF-42, in that: (1) The dispersion of item difficulties within each subscale sufficiently covered each content domain, (2) the dispersion among aggregate subscale difficulties within the overall K-SF-42 scale also showed a satisfactory degree of coverage, and (3) the rank ordering of subscale difficulties across samples showed a relatively high degree of cross-cultural invariance.
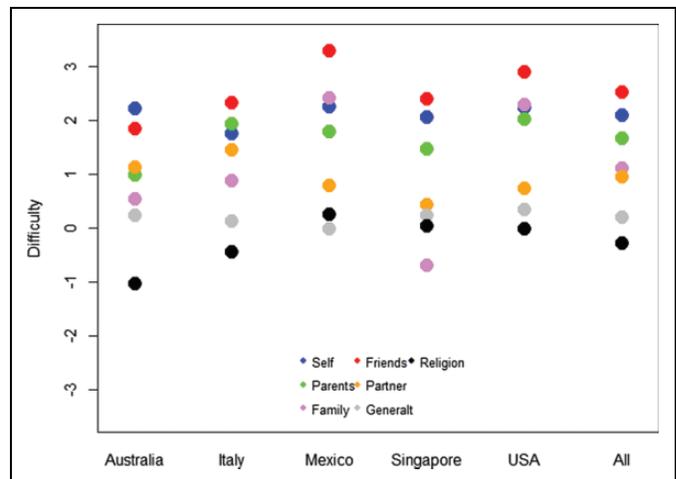
Figure 1. Item response theory subscale difficulties for K-SF-42.

Appendix

The K-SF-42.

Please indicate how strongly you agree or disagree with the following statements. Use the scale below and write your answers in the spaces provided. For any item that does not apply to you, please enter "0."

Disagree Strongly	Disagree Somewhat	Disagree Slightly	Don't Know/Not Applicable	Agree Slightly	Agree Somewhat	Agree Strongly
−3	−2	−1	0	+1	+2	+3

1._____When faced with a bad situation, I do what I can to change it for the better
2._____When I encounter problems, I don't give up until I solve them.
3._____I find I usually learn something meaningful from a difficult situation.
4._____When I am faced with a bad situation, it helps to find a different way of looking at things.
5._____Even when everything seems to be going wrong, I can usually find a bright side to the situation.
6._____I can find something positive even in the worst situations.
7._____I spend a great deal of time per month giving informal emotional support to my blood relatives.
8._____I contribute a great deal to the welfare and well-being of my blood relatives in the present.
9._____I spend a great deal of time per month giving informal emotional support to casual acquaintances (such as neighbors or people at church).
10._____I contribute a great deal to the welfare and well-being of my friends these days.
11._____I spend a great deal of time per month doing formal volunteer work at school or other youth-related institution.
12._____ I often contribute to any other organizations, causes, or charities (including donations made through monthly payroll deductions).
13._____ I'm a very religious person.
14._____ Religion is important in my life.
15._____ Spirituality is important in my life.
16._____ I closely identify with being a member of my religious group.
17._____ I frequently attend religious or spiritual services.
18._____ When I have decisions to make in my daily life, I often ask myself what my religious or spiritual beliefs suggest I should do.
19._____ I worry that romantic partners won't care about me as much as I care about them.
20._____ I don't feel comfortable opening up to romantic partners.
21._____ I want to get close to my partner, but I keep pulling back.
22._____ I often want to merge completely with romantic partners, and this sometimes scares them away.
23._____ I am nervous when partners get too close to me.
24_____ I find that my partner(s) don't want to get as close as I would like.

The following are some questions about means of help that people offer each other. Use the scale below and write your answers in the spaces provided, indicating about how often any parent, family member, or friend has helped you in each of the following ways. For any item that does not apply to you, please enter "0."

Not At All	A Little	Some	A Lot
0	1	2	3

While you were growing up…

25._____ How much time and attention did your biological mother give you when you needed it?
26._____ How much effort did your biological mother put into watching over you and making sure you had a good upbringing?
27._____ How much did your biological mother teach you about life?
28._____ How much love and affection did your biological father give you while you were growing up??
29._____ How much time and attention did your biological father give you when you needed it?
30._____ How much did your biological father teach you about life?

During the last month…
31._____ How much have your relatives helped you get worries off your mind?
32._____ How much have your relatives told you that you had done something well?
33._____ How much have your relatives told you that they liked the way you are?
34._____ How much have your relatives shown you affection?
35._____ How much have your relatives listened to you when you talked about your feelings?
36._____ How much have your relatives shown interest and concern for your well-being?
37._____ How much have your friends helped you get worries off your mind?
38._____ How much have your friends told you that you had done something well?
39._____ How much have your friends told you that they liked the way you are?
40._____ How much have your friends shown you affection?
41._____ How much have your friends offered to take you somewhere?
42._____ How much have your friends shown interest and concern for your well-being?

References

Abed, R., Mehta, S., Figueredo, A. J., Aldridge, S., Balson, H., Meyer, C., & 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

Acar, T. (2011). Sample size in differential item functioning: An application of hierarchical linear modeling. *Educational Sciences: Theory & Practice*, *11*, 284–288.

Alonso-Arbiol, I., Balluerka, N., & Shaver, P. R. (2007). A Spanish version of the Experiences in Close Relationships (ECR) adult attachment questionnaire. *Personal Relationships*, *14*, 45–63.

Barrera, M. Jr., Sandler, I. N., & Ramsay, T. B. (1981). Preliminary development of a scale of social support: Studies on college students. *American Journal of Community Psychology*, *9*, 435–447.

Bond, T., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). New York, NY: Routledge.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum.

Brennan, K. A., Clark, C. L., & Shaver, P. R. (1998). Self-report measurement of adult attachment: An integrative overview. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 46–76). New York, NY: Guilford Press.

Brim, O. G., Baltes, P. B., Bumpass, L. L., Cleary, P. D., Featherman, D. L., Hazzard, W. R., . . . Shweder, R. A. (2000). *National survey of midlife development in the United States (MIDUS), 1995–1996* [Computer file]. ICPSR version. Ann Arbor, MI: DataStat, Inc./ Boston, MA: Harvard Medical School, Department of Health Care Policy [producers], 1996. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Buunk, A. P., & 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., & 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., & 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., & 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.

Egan, V., Figueredo, A. J., Wolf, P., McBride, K., Sefcek, J. A., Vásquez, G., & Charles, K. (2005). Sensational interests, mating effort, and personality: Evidence for cross-cultural validity. *Journal of Individual Differences*, *26*, 11–19.

Figueredo, A. J., Andrzejczak, D. J., Jones, D. J., Smith-Castro, V., & 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., Cabeza de Baca, T., Black, C. J., Garcia, R. A., Fernandes, H. B. F., & Wolf, P. S. A., & Woodley of Menie, M. A. (2015). Methodologically sound: Evaluating the psychometric approach to the assessment of human life history [Reply to Copping, Campbell, and Muncer, 2014]. *Evolutionary Psychology*, *13*, 299–338.

Figueredo, A. J., Cabeza de Baca, T., & 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-Ruíz, 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., Jacobs, W. J., Gladden, P. R., Bianchi, J. M., Patch, E. A., Beck, C. J. A., . . . Jiang, Y (2016). *Intimate partner violence, interpersonal aggression, and life history strategy.* Manuscript submitted for publication.

Figueredo, A. J., & Olderbak, S. G. (2008). Generalizability theory analysis for psycholinguistic applications. *Journal of Memory and Language, 59,* 426–433.

Figueredo, A. J., & 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., & Schneider, S. M. R. (2004). The heritability of life history strategy: The K-factor, covitality, and personality. *Social Biology, 51,* 121–143.

Figueredo, A. J., Vásquez, G., Brumbach, B. H., & Schneider, S. M. R. (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., Vásquez, G., Brumbach, B. H., Sefcek, J. A., Kirsner, B. R., & Jacobs, W. J. (2005). The K-Factor: Individual differences in life history strategy. *Personality and Individual Differences, 39,* 1349–1360.

Figueredo, A. J., & 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.

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 & 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., & Corral-Verdugo, V. (2010). Executive functions, attention and juvenile delinquency. In M. Frias-Armenta & V. Corral-Verdugo (Eds.), *Bio-psycho-social perspectives on interpersonal violence* (pp. 29–50). Hauppauge, NY: Nova Science.

Garcia, R. A., Cabeza de Baca, T., 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,* 23–43.

Gaxiola-Romero, J., & 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.

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

Linacre, J. M. (2013). Differential item functioning DIF sample size nomogram. *Rasch Measurement Transactions, 26,* 1391.

Linacre, J. M. (2016). *Winsteps® Rasch measurement computer program.* Beaverton, OR: Winsteps.com.

Microsoft. (2003). *Microsoft Excel* [computer software]. Redmond, Washington: Microsoft.

SAS Institute Inc. (2011). *Base SAS® 9.3 Procedures Guide: Statistical Procedures.* Cary, NC: SAS Institute Inc.

Mitchell, T. R. (1982). Motivation: New directions for theory, research, and practice. *Academy of Management Review, 7,* 80–88.

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.

Patch, E. A., Figueredo, A. J., Garcia, R. A., & Kavanagh, P. S. (2016) *E pluribus unum: Rethinking the Dark Triad.* Manuscript submitted for publication.

Scott, N. W., Fayers, P. M., Aaronson, N. K., Bottomley, A., de Graeff, A., Groenvold, M., . . . Sprangers, M. A. (2009). A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales. *Journal of Clinical Epidemiology, 62,* 288–295.

Soothill, W. E. (1937). *The analects of Confucius.* London, England: Oxford University Press. (Original work published 1910)

Sotomayor-Peterson, M., Cabeza de Baca, T., Figueredo, A. J., & 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., & Taylor, A. R. (2012). Couples' cultural values, shared parenting, and family emotional climate within Mexican-American families. *Family Process, 51,* 218–233.

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

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

Tristan, A. (2006). An adjustment for sample size in DIF analysis. *Rasch Measurement Transactions, 20,* 1070–1071.

van der Linden, D., Figueredo, A. J., de Leeuw, R. N. H., Scholte, R. H. J., & 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.

Woodley of Menie, M. A., & 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.